

PT	REPEAT	2778	2781	6.	Db	2202	DTWQVLTNEHEFTKQVSNEPTMFDNSKNSBET-LTKAVTDSKTIISLU-KGVITE	2257
PT	REPEAT	2782	2785	7.	SQ	2876 AA;	331433 MW;	6B7D8CA71AFBF0D3 CRC64;
Query	Match	4.3%	18.9%	Score 231; DB 1; Length 287;	Matches	243; Conservative 219; Mismatches 436; Indels 388; Gaps 65;		
DB	Best Local Similarity				QY	11 LALTTWVWTTSQEVYGLERERESVKEEQTQS---ASEDWEEDEERKVNKSSENIV- 65		
QY	1314 IATKKKGTSQDSQDINEL--ESTIKEVHNQLQVQESNSMEE--MRKQISMDLIL 1368				Db	2258 VNEVENTNTTESSAKETIAALYNS- 66		
QY	6 --DETIVSDFPSDG-----NSNNSSKTESVVS----DPKVPKAKPETOQASW 108				QY	844 KSGKLUVALDMSDTIGEGORDAVCNPLINVDNEGYHTLAVATLADVEGLYKDLINSS 903		
Db	1369 NNSETTAKETISNNNTQVALGERENAKTUNLQVAMTEAKAHKNTIDALEDQ 1428				Db	2313 VS-----KIFNTVLDTOK--SNIVTNQHSINN-----VKDKLKGK 2345		
QY	10 9 SSNDASKVEVPRQDTASKETLETSTWAKDFTVTRGDTLVGFSKGINKUSQTSHLVLP 168				QY	904 LDK-TKAIRQIPLAKYHRIG-IFQAIRAAEADRLLPKPKGSYNEVNYRKQCEKML 961		
Db	1429 IDTEVSKIEQINRETMANKOBIKSYISEKEYKDCTETISNSKRGKDKIEPLE---- 1482				Db	2346 LQELIDADSSFTLESIIKKFNETVSHKTNIGELFQ-LOOTNKSEHDNTVARKKEKIVH-- 2401		
QY	16 9 HAADGTSQLTQVAFASAPTPDK----KTAIABYTSRIGENKPSRLID--QKEIIDEGETI 221				QY	962 KPVYKTPPNKALPNEKUDGDRAGKH-----NINAEATNSVAVTPIRSE 1007		
Db	1483 -----KFKPNEESENNSNKVNINBENENI-RNSEQYQLKIEDAEKQASTKVEL 1527				Db	2402 -----LIRNV----ESIKGD--VKNHDDQYMKKLNASLNUINDNQVNTNSINISDEELK 2448		
QY	222 FNAQYQTKLTIPNGYKSIQGDAFVUNKNIAEVNLPSELETSDYFAHMSLKVQVLPDNU 281				QY	1008 QQLHKSQSDVNLPQTSKSKNFVITL 1033		
Db	1528 FHKHE----TTISNIFKE-SEITGVETKSQKKKABDINGKBERHNEIOTQVKQFOENL 1583				Db	2449 KLIKVEENDQLCNNNTQFISDM 2474		
QY	282 KVIGELAFTPQNOIGGKYLPRHLIKLARAFKSNRQITVERLGSKUKVIGASFOD-MN 339							
Db	1584 NKLNEPHNYDN-----ADELNNDK-----STNAKULIETNLESVKN 1621							
QY	340 LRNVM-LPDGLKETES-----EAFTGNGDDE-----HYNQNVVLRTR 376							
Db	1622 ISEITNIKKQGKBEKISKAKDQMOKIKATSENATAEKLEKVKQDDQSNVNVNQIT--T 1677							
QY	377 GQNPHQOLATE-----NTVNPDSKL--WRATPMDYTKLLE-----DFTYQK 417							
Db	1678 ERN-LIVTEKNRNLGIDSTTNTNIEGALKESKGNYEVIGFLEEGKRNKLKVDTIKS 1735							
QY	418 -NSVTG---FSNKGQLQVKVRNKL-----ELPKQHNGITTEIGNAFRV--- 459							
Db	1736 INSTVGNFSSLNFNDQYDFNKNINDYENKNGEYNEFEG-SLINKSEN-LRNASENT 1793							
QY	460 -DFQS-KTLLR-----KYL-----BTKLPLSTIRKIGAFARQFQSNNIKSFAESEDELI- 505							
Db	1794 SDYNSAXTIRLRAQKERVNLLNKEEBAKVKYLDVKVKEFPR-----IFNMKESLDKIN 1847							
QY	506 ---KEGRFMNHRIGT-DLKQKLKIGDAFHINIIYATVLP-----ESTOEGRSFA 555							
Db	1848 EMIKKEQLTIVNEGHGNVQVQLVENIKELWDE---NNISLQKATQGKNEEIQKITHSTLK 1903							
QY	556 QNGALHUMFIGMKVKTG---EMARLNSKLESVNLSEQKOLK-----TIEQAFSD 603							
Db	1904 NRKATLGHVDSIAYVYKTPBELA-LTELIGDAKLTQABLKFKESKNVYLETENMSK 1962							
QY	604 NALSEVUPLPNLQ-----TIREFAK-----RNHL 628							
QY	1963 NT-NELDWTNQNTDAYKVALETLAHSDEBDTQKQDSKLIENGNOIYKVULINQYKNI 2021							
Db	629 KEYRGS-----STLSQTF-----NAFDQ---N 648							
QY	2082 NSDSKLERIKTDPSLKNALKTIEGEVNAKASSDNHHEVJSKSEPVNPALSEIEKETD 2141							
Db	2022 SSIKSKEBAVKIGNVSKKHSBLSKTCDSKSYDNTIALEKQTELONLRSNSTQEKINT 2081							
QY	649 DGDKRFKG-----KVVRT-----HNNSHM-----LADGEPRID 678							
Db	2082 NSDSKLERIKTDPSLKNALKTIEGEVNAKASSDNHHEVJSKSEPVNPALSEIEKETD 2141							
QY	679 POKLUSSTMDIERTKUMLIKEGLYUPLQTTQFQPREMTA-----GKALLS--KSNLR 729							
Db	2142 IDSLNTADLLERKGRTEVSKYKLKOTVKEISDPTLINTIEKVKAYLIVKNEYE 2201							
QY	730 QGKQKDFQBAQFQFLGV-----DLDKAIKAKBALKVTKATRNGHLLERSINKAVL 783							

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on:

April 27, 2004, 11:10:26 ; (without alignments)

3231.413 Million cell updates/sec

Title: US-10-091-007A-24

Perfect score: 5352

Sequence: 1 MTKKHLKTLALALTTVSVVT..... VSLCLLFLVTAGKKRARK 1055

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.0

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	231	4.3	2867	1	RBP2_PLAVB	Q00799 plasmidium
2	224	4.2	2748	1	NUM1_YEAST	Q00402 saccharomyces
3	215.5	4.0	1727	1	ALMI_SCIRBO	O9utk5 schizosaccharomyces
4	209	3.9	1208	1	FCP1_SCIRBO	Q92351 schizosaccharomyces
5	209	3.9	1875	1	MPL1_YEAST	Q02455 saccharomyces
6	208	3.9	1013	1	SCA4_RICKH	O9aj81 rickettsia
7	203	3.8	998	1	SCA4_RICKH	O9aj99 rickettsia
8	202.5	3.8	2869	1	RBPL1_PLAVB	Q00798 plasmidium
9	202	3.8	1022	1	SCA4_RICKH	O9zd49 rickettsia
10	200	3.7	872	1	TBP5_TETTH	Q94819 tetrahymena
11	200	3.7	939	1	IF2_WOLSU	Q7mtx5 wolinella
12	199.5	3.7	3259	1	G001_HUMAN	O14789 homo sapien
13	199	3.7	1357	1	SP001_SCIRBO	Q86up2 homo sapien
14	199	3.7	1957	1	SP002_SCIRBO	Q01411 schizosaccharomyces
15	198.5	3.7	1790	1	US001_YEAST	Q25386 saccharomyces
16	198	3.7	1010	1	SCA4_RICKH	O9aj75 rickettsia
17	196.5	3.7	992	1	SCA4_RICKH	O9aj63 rickettsia
18	195.5	3.7	1969	1	MYZA_CABEL	P12844 caenorhabditis
19	192.5	3.6	1358	1	ST4_YEAST	P11978 saccharomyces
20	192.5	3.6	1839	1	CTIA_SACKL	Q10411 schizosaccharomyces
21	192	3.6	1022	1	SCA4_RICKH	Q05268 rickettsia
22	189.5	3.5	981	1	SCA4_RICKH	O9aj37 rickettsia
23	188.5	3.5	1012	1	SCA4_RICKH	O9aj80 rickettsia
24	188	3.5	1011	1	SCA4_RICKH	O9aj83 rickettsia
25	187	3.5	1928	1	MY001_YEAST	Q08964 saccharomyces
26	186.5	3.5	1011	1	SCA4_RICKH	O9aj82 rickettsia
27	185	3.5	1018	1	SCA4_RICKH	O9aj79 rickettsia
28	185	3.5	1012	1	MY002_RABIT	Q53748 orctotilus
29	184.5	3.4	1539	1	Y273_HOMTA	O15078 homo sapien
30	183.5	3.4	171	1	CHAO_DROME	Q12024 drosophila
31	183.5	3.4	8545	1	ANCI_CABEL	O94m44 caenorhabditis
32	183	3.4	2663	1	CENE_HUMAN	Q02244 homo sapien
33	181	3.4	991	1	SCA4_RICKH	O9aj77 rickettsia

#### ALIGNMENTS

34	181	3.4	1324	1	SMC4_SCIRBO	P41004 schizosaccharomyces
35	181	3.4	1690	1	C190_DROME	Q8vje5 drosophila
36	180.5	3.4	1979	1	TRIA_HUMAN	Q5f643 homo sapien
37	179.5	3.4	1163	1	SBCC_CLOAB	Q97fk1 clostridium
38	179.5	3.4	1972	1	MYB_HUMAN	P35749 homo sapien
39	179	3.3	1312	1	RA50_YEAST	P12753 baccharomyces
40	179	3.3	3210	1	CENF_HUMAN	P89454 homo sapien
41	178.5	3.3	1159	1	EXB_BORBU	Q51578 borrelia buchneri
42	178.5	3.3	1658	1	YM67_YEAST	Q03661 saccharomyces
43	177.5	3.3	1701	1	MSPI_PLAFF	P13819 plasmidium
44	177.5	3.3	1972	1	MYB_MOUSE	Q08638 mus musculus
45	177	3.3	1637	1	MRSP_STAU	P80544 staphylococcus

RA Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,  
 RA Du Z., Pavello A., Fulton L., Gattung S., Gracco T., Hallsworth K.,  
 RA Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L.,  
 RA Kirtsen J., Kocab T., Langston Y., Latreille P., Le T., Mardis E.,  
 RA Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifkin L.,  
 RA Riles L., Taich A., Trevaskis E., Vignati D., Wilcox L., Wohldman P.,  
 RA Vaudin M., Wilson R., Waterston R., Albermann K., Hani J., Heumann K.,  
 RA Kleine K., Mewes H.W., Zoller A., Zaccaria P.;  
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV.";  
 RL Nature 387:78-1987.  
 CC -!- FUNCTION: Controls nuclear migration. NUM1 specifically controls  
 the interaction of the bud neck cytoskeleton with the pre-  
 divisional G2 nucleus perhaps by recognizing G2-specific  
 cytoplasmic microtubuli or other components of the nuclear  
 envelope.  
 CC -!- MISCELLANEOUS: Additional regions of lower homology to the repeat  
 consensus (always starting with proline) are found in both  
 flanking domains of the tandem repeats.  
 CC -!- SIMILARITY: Contains 1 PH domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 use by non-profit institutions. There are no restrictions on its  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [licensee@isb-sib.ch](mailto:licensee@isb-sib.ch)).

CC  
 EMBL: X61236; CAA43554\_1; -;  
 DR Germoline; Z50046; CAA90372\_1; -;  
 DR SGD; S002557; NTM1; -;  
 DR GO; GO:0005934; C: bud tip; IDA.  
 DR GO; GO:0005938; C: cell cortex; IDA.  
 DR GO; GO:0015631; F: tubulin binding; IDA.  
 DR GO; GO:000065; P: microtubule cytoskeleton organization and b. . . ; IMP.  
 DR InterPro; IPR001839; DUF321.  
 DR InterPro; IPR001839; PH.  
 DR Pfam; PF03778; DUF321; 13.  
 DR Pfam; PF00169; PH; 1.  
 DR SMART; SM0223; PH; 1.  
 DR PROSITE; PS5003; PH\_DOMAIN; 1.  
 KW REPEAT.  
 FT DOMAIN 593 1384 12.5 X TANDEM REPEATS.  
 FT REPEAT 593 656 1.  
 FT REPEAT 657 727 2.  
 FT REPEAT 728 798 3.  
 FT REPEAT 799 862 4.  
 FT REPEAT 863 926 5.  
 FT REPEAT 927 990 6.  
 FT REPEAT 991 1054 7.  
 FT REPEAT 1055 1118 8.  
 FT REPEAT 1119 1182 9.  
 FT REPEAT 1183 1246 10.  
 FT REPEAT 1247 1310 11.  
 FT REPEAT 1311 1374 12.  
 FT REPEAT 1375 1384 13. (INCOMPLETE).  
 FT DOMAIN 2573 2683 PH.  
 FT CONFLICT 1570 1570 A -> V (IN REF. 1).  
 FT CONFLICT 1822 1822 E -> K (IN REF. 1).  
 FT CONFLICT 1862 1962 KAS -> RHL (IN REF. 1).  
 FT CONFLICT 1971 1972 KD -> RN (IN REF. 1).  
 FT CONFLICT 2049 2049 S -> N (IN REF. 1).  
 FT CONFLICT 2637 2637 V -> A (IN REF. 1).  
 SQ SEQUENCE 2748 AA; 313030 MW; EB4B48PF95021142 CRC64;  
 Query Match 4.2%; Score 224; DB 1; Length 2748;  
 Best Local Similarity 20.0%; Pred. No. 0.011; Matches 228; Conservative 191; Mismatches 464; Indels 258; Caps 50;  
 QY 7 KTLALALTIVSVVWVTSQEVYGLERESTKQEQTOSASEDDWFEEDNERKTKTNSKENSTD 66

Db 193 ESLKLKLTDLRILTILTKN-GMENDS---SOKHYDKRISELKERIDLNNDRLL 245  
 QY 67 BTVSILPSDGNSNNNSKTESVSPDKQVRKAKPETO-----EASNNSNASKVE 117  
 Db 246 ISVSLTSEINS-LQSNRER-IKQKQLDKAKASISLKRKVQKYYQKHTSTVT 303  
 QY 118 VPKQPTASKETLETSTWEAKDFVTRGDTLVGFSPSGINKLQSQSHLVLPSHAAGTQLT 177  
 Db 304 DPDSIGTTSBEDIIVIDHMTGTPSVEDISBDLKVYSEKRNMLNLSDSYRNLLQ 363  
 QY 178 QVASAAFTPKKTAIAVTSRIGENGKPSRSLDQKKEIDBEGFFNAYQITKUTFPNGY 237  
 Db 364 KSEASKPKDODELMYKEVANTNMALPNDNYSKESFESLH-----KYLEASGY 416  
 QY 238 SIGDQFVDNKNAIAVN---LPESLETT----- 262  
 Db 417 VLPPLERF-ENLNESNSNPSPSYNLKEQGQALKKIPIDQSTPNLUKEPTIFPLPLTSKIDC 475  
 QY 263 -----SDY-----APAHMSKQVK-----LPDNL-----KVIGELAFFF-D-N 292  
 Db 476 LIIPKQDINDLFESPVKNPSIEOMKCLEAKNDLQNSICKWLEERNGCKWLNLDLXFSMVN 535  
 QY 293 QI-GGKLUPLRHI-----IKLAFRAFKNRQTEVFQGSKLKVGEASFQDNLU-RNV 343  
 Db 536 KIETPSQVQYSLDKAKYDQVLTIDTAKLEGKPNPTDFREK-----ASDYLILKKED 589  
 QY 344 MLPQGKLEKSEATGNP---GDEHYNQ-----VLRFTGQPHOLATENTYV 391  
 Db 590 YVSPSPLEYLVEHAKTNHHLSDSAYEDLVKCKENPDMFLKERSAKOHTVSNHAYVE 649  
 QY 392 PDKSLIWRATPDMDYKWLIEDFTYQKNSVTGFSNKGLOKVRNKNLNEIPKHOHGTITI 451  
 Db 650 LEKLK---RQPSLEVY---VEHAKATNHLSDSAYEDLVKCKENPDMFLKERS---AKL 701  
 QY 452 GDNAAFRNVDFOSKTIRKY-DLE-EIKSTIRKIGAFAFQSNLKSPESEDIEKIGA 509  
 Db 702 GHTVWSN-BAYSELQKYSLEKEVEQPSLAVLVEHAKATNHLSDSAYEDLVKCKENP 760  
 QY 510 FMNHRIGTIDLKDLKLIKQDAFHINHIIYAVLPESVQEVIGRAFPRONGALHMFIGNKV 569  
 Db 761 DVE-----FLKEKSAKG-----HTVWSSEBEYSLQR----- 787  
 QY 570 KTIGEMAFISNKCLESSVNISE-QQKLTTEVOQAFSDNALSEVVL---PPNLQITREAFR 625  
 Db 788 ---KYSLEKEVEQPSLAVLVEHAKATNHLSDSAYEDLVKCKENPDMFLKERSKL 843  
 QY 626 NHLKVKGSSTLSQITENAFDQNDGKQPKGKKYVWRTHINNSMLADGERFILLPPKUST 685  
 Db 844 GH-----TVVSNEAYSELEKKLQPSLAVLVEHAKATNHLSDSAYEDLVKCKENP 896  
 QY 686 MVDL-EKVIK---IIEGDYSTL-RQTTQTREMITAGKA---LISKS---NLR 729  
 Db 897 VEFLEKESAKLGLHTVVSNBAYSELEKKLQPSLAVLVEHAKATNHLSDSAYEDLVKCK 956  
 QY 730 QGKKQKFLQEAQFLGR-VLDLGAIKAKKA-----LVTKKATKQHIL----- 772  
 Db 957 ENPDMEFLKEKSALKHTVVSNBAYSELEKKLQPSLAVLVEHAKATNHLSDA 1016  
 QY 773 -----ERISKAVLAVNSAIKKANVKRKEKELDILTDLVEGKGP--LAQAT 817  
 Db 1017 VCKKRNPDMEFLKEKSALKHTVVSNBAYSELEKKLQPSLAVLVEHAKATNHLSDA 1076  
 QY 818 MVQCVVLUKTPPLPEYYIGLNNTF--DKSGKLUYALMSDITGEGERGDAYNPLN--V 873  
 Db 1077 YEVLKCKENP-----DVEFLKESAKLIGHTVVSNEAYSELEK-KLEQPSLAVL 1125  
 QY 874 DEDEGYHTLAVATLAM---YEGYIKDILSSDKTVAIRQPLAKHRLGFOAIRNA 931  
 Db 1126 EHAKATNHL---LSDSAVEEL-VCKKENPDMFLK-----EKSALKHTVVSNEY 1173  
 QY 932 AEDRDLPKTPKGYLNENPVNKKQKOMKOLVKDVYKPIPNKALPNEKGDKRAAK-GHN 990  
 Db 1174 SELEKUKLQPSLAVLVB---HAKATDHLSDSAYEDLVKCKENPDPVFLKEKSALKHT 1230

Or 991 I 991 CC  
 Db 1231 V 1231 CC

RESULT 3  
 ALM1\_SCHPO STANDARD PRT; 1727 AA.  
 ID ALM1\_SCHPO  
 AC Q9UTKS; O13313; Q9UTR8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Abnormal long morphology protein 1 (Sp8).  
 GN ALM1 OR SPAC1486.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycotae; Schizosaccharomycetidae; Schizosaccharomyces.  
 OC NCBI\_TaxID:4896;  
 RN [1] SEQUENCE FROM N.A.  
 RP  
 RC STRAIN=97;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver P., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Sketton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoperez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Reinhardt R., Rohl T.M.,  
 RA Borzny K., Langer T., Beck A., Lehrh H., Reinhardt R., Rohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Goux S., Lelaique V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochat M., Gaillardin C., Tallada V.A., Garzon A., Mode G.,  
 RA Daga R.R., Cruzad L., Jimenez J., Sanchez M., del Rey F., Bentito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,  
 RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;  
 RT The genome sequence of Schizosaccharomyces pombe.;  
 RL Nature 415:871-880(2002).  
 RN [2]  
 RP SEQUENCE OF 495-1727 FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=97;  
 RX MEDLINE=20123449; PubMed=10560053;  
 RA "The alm1+ gene from Schizosaccharomyces pombe encodes a coiled-coil  
 protein that associates with the medial region during mitosis.";  
 RL Mol. Gen. Genet. 262:921-930(2000).  
 RN [3]  
 RP SEQUENCE OF 644-834 FROM N.A.  
 RC STRAIN=968 b90;  
 RX MEDLINE=2022368; PubMed=10759889;  
 RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,  
 RA Hirakawa Y.;  
 RT "Large-scale screening of intracellular protein localization in living  
 yeast cells by the use of a GFP-fusion genomic DNA library.";  
 RL Genes Cells 5:169-190(2000).  
 CC -!- FUNCTION: AFTER THE ONSET OF MITOSIS, AT MID- TO LATE ANAPHASE,  
 CC CO-LOCALIZES WITH THE MEDIAL ACTIN RING. MAY PLAY A ROLE IN  
 CC CYTOKINESIS.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 DR EMBL; AL133357; CAB62414; 1; -; -  
 DR EMBL; AF010473; AAB65416; 1; ALT\_INIT.  
 DR PIR; T50073; T50073.  
 DR GenBank\_Spombe; SPAC1486.04C; -.  
 DR Coiled coil.  
 FT DOMAIN 57 361 COILED COIL (POTENTIAL).  
 FT DOMAIN 443 463 COILED COIL (POTENTIAL).  
 FT DOMAIN 442 740 COILED COIL (POTENTIAL).  
 FT DOMAIN 804 1106 COILED COIL (POTENTIAL).  
 FT DOMAIN 1223 1427 COILED COIL (POTENTIAL).  
 FT DOMAIN 1497 1555 COILED COIL (POTENTIAL).  
 FT DOMAIN 1601 1664 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 1727 AA; 197858 MM; P820BR8D9C132644 CRC64;  
 Query Match 4 0%; Score 215 5; DB 1; Length 1727;  
 Best Local Similarity 19.3%; Pred. No. 0.015; Gaps 57;  
 Matches 231; Conservative 198; Mismatches 422; Indels 343; Gaps 57;  
 QY 12 ALTTTSWVITQSQEVGLERESVKQEQTOSASED----DWFEEDNERKINSKSTD 66  
 DB 116 ALTT-----ENGIRRANSELQEQSKIASBOSIAKQDQIAQLEAQNENSHLGQVOSAH 167  
 QY 67 BTVSD-----LESQDGNNNSKTESVSPDKQPKPEVTEASNSNSPDRSKVE 118  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 223 NAVQITK-----LTTPNGYKSGQ--DAFVNKNIAEVN-----LPE 257  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 308 AERAFKNSRNLQTVERLIGSKLKVIGEAFODNNLNNMPDPGLEKESAAFTGNGPDEH 367  
 DB 431 QRSFVRRNYY-MNFSSESLET----SNNNLTKVQABLLSTKMRQEA-----YL 475  
 QY 368 NQVWVTRGQNGPQ---LATENWVNDKSLRATGPDMYKWLIEDPTYQNSVIG-- 422  
 DB 476 QLTASRTOQCDSDLSRIVCILMAELDHLNETS---RNVNPATVQVOLA-LDE---YAQNPSTAE 529  
 QY 423 -FSNKGLO---KURNKMLEPQHNGNITITRIGDIAFRNVDQSKTURKIDLEETKL 476  
 DB 530 TLMNKELMFSITKEAVSKTLEUREKVALEC-----DVEIQKQTVQKQISNAVKE 580  
 QY 477 PSTRKIGAFAFQSNLNKFEASDLEETKEGAFMNR-----IG 516  
 DB 581 NSNT-----LSEQKMLESELMNSKKNESLNLNLKEMATRSRSLSHNSAG 632  
 QY 517 TLDKDKLKLIGDAFHNTMIVALPPE-SVQEIGRAFRQNGALHMFQIGKVKTIG 574  
 DB 633 NID-DROMKSIDESTRELFNEVYRNEMTAO---SLSKRNDL-----LSE 676  
 QY 575 MAFLSKLNKESVNLSEQKQKLT-----IEVQAFSDNALESEVLPNPLTI--REAFKR 625  
 DB 677 MEAIRKELE--NSKVQQQSLDRUANNDVEAFKEAKLRSINQNLQDITSRQDQRAS 734







us-10-091-007a-24.rsp

QY 840 VYFDKSGKLUYALDMSTI---GERGOKUAYGNPLNIVNE 875  
Db 909 LIDDKNLSLKVRLOTIDKLUQEQOTOKQAIENPSVKTD 948

RESULT 7

SCA4\_RICKAK STANDARD; PRT; 998 AA.

ID SCA4\_RICKAK STANDARD; PRT; 998 AA.

AC Q9AIX9; 16-OCT-2001 (Rel. 40, created)  
DT 16-OCT-2001 (Rel. 40, last sequence update)  
DT 16-OCT-2001 (Rel. 40, last annotation update)  
DR Antigenic heat-stable 120 kDa protein (PS120) (120 kDa antigen)  
DE (Protein PS 120) (Fragment).  
GN SCA4 OR D.

OS Rickettsia akari.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Rickettsiaceae; Rickettsia.

OX NCBI\_TaxID=786;

RN [1] RP SEQUENCE FROM N.A.

RA Sekeyova Z., Roux V., Raoult D.,  
RT "phylogenetic analysis of Rickettsia spp. by comparing sequence of the  
RT gene D, coding for an intracytoplasmic protein.",  
RT submitted (DEC-1999) to the EMBL/GenBank/DBTG databases.  
CC !- SUBCELLULAR LOCATION: Cytoplasmic (probable).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are NO restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to licensee@isb-sib.ch).

DR EMBL; AF213016; AAK0691.1; -.

KW Antigen.

FT 1 1  
FT NON\_TER 998 998 MW; FCB843AC62DE5BD5 CRC64;

SQ SEQUENCE 998 AA; 109328 MW; FCB843AC62DE5BD5 CRC64;

Query Match 3.8%; Score 203; DB 1; Length 998;  
Best Local Similarity 18.9%;  
Matches 220; Conservative 176; Mismatches 405; Indels 360; Gap 55;

QY 6 LKTLALALITVAVVTSQEVQVGLEREEVKQEQTQASE-----DDFEEDNERKT 56  
Db 57 ISLISGPISTDQI--SDPITKAVREITIQQQDEAQLKDALAVIDRQLAOKRKEI 114  
Qy 57 NYSKNSTVDETDVSDLNSDGNSNSSSKTESVSDPKQVKAKPE-----VT 103  
Db 115 EEEKEK--DKKLUSVFF--GNPAN---REFDIALEKPELKKKLESETIGYKMLT 164

QY 104 QDASN-----SENDASKEVKPDKTAKETSTWEEKAFVTRDILV 148  
Db 165 YSANGYHGGFKEPVQWENQISASDL-RATVVKNDAGDELCTINETTVTKPFTVA--- 218

QY 149 GISKSG-----INKLQSTSHVLPSHADQVLTQVASFAFTPDKTAT 192  
Db 219 -- KRDGTQVQINSYRAIDFPPIKULKADQMSHMSWALKADGPK-----PSKDRAV 266

QY 193 --AEYTSRIGENGKPSRLDIDKEIIDEGETFNAYQOLTKLITPQNGYKSIGDAFWDNKN 249  
Db 267 YFTAHYEE--GPNGKP----- 280

QY 250 TAEVNLPLPSELETISDYAHFAMSILQVKUDPNLKVIGE---LAFFDNQIGGKLY--LPR 302

RESULT 8

RPB1\_PLAVB STANDARD; PRT; 2869 AA.

ID RPB1\_PLAVB STANDARD; PRT; 2869 AA.

AC Q00798; 01-APR-1993 (Rel. 25, created)  
DT 01-APR-1993 (Rel. 25, last sequence update)  
DT 10-OCT-2003 (Rel. 42, last annotation update)  
DR Reticulocyte binding protein 1 precursor.

GN RBP1.

OS Plasmodium vivax (strain Belem).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=31273;

QY [1] RP SEQUENCE FROM N.A.  
RX MEDLINE=92315338; PubMed=1617731;  
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
RT "A reticulocyte-binding protein complex of Plasmodium vivax  
membranes";  
RL Cell 69:1213-1226 (1992);  
- FUNCTION: involved in reticulocyte adhesion. Specifically binds to

CC - human reticulocyte cells.  
 CC -!- SUBUNIT: Homodimer (Potential).  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC

CC DR EMBL; MP8097; AAA29743.1; -

CC KW Malaria; Receptor; Signal; Transmembrane.

CC FT SIGNAL 1 17; POTENTIAL.

CC FT CHAIN 18 2869; RETICULOYTE BINDING PROTEIN 1.

CC FT DOMAIN 18 2807; EXTRACELLULAR.

CC FT TRANSMEM 2808 2826; CYTOPLASMIC.

CC FT DOMAIN 2827 2869; CELL ATTACHMENT SITE (POTENTIAL).

CC FT SITE 1030 1032; CELL ATTACHMENT SITE (POTENTIAL).

CC FT SEQUENCE 2869 AA; 330213 NW; B9DB442205BCFF CRC64;

CC

Query, Match 3-8%; Score 202.5; DB 1; Length 2869;  
 Best Local Similarity 17.9%; Pred. No. 0.12; Mismatches 211; Conservative 178; Indels 385; Gaps 50; Matches 211;

QY 4 KHLKHLALLTT-VSVVTSQEVIGLERESVKQEQTOSASDDWF---RED--- 51

QY 52 -----NERKTNTSKENSTUDTVDSDLTSDGNANNSSKTESVSDPKQVPAK- 886

QY 887 EBQEVTKHKNFSRKRGEKSAEITMREVKIESQNLNTYVIEKFSLIGDNEVSTAK 946

QY 99 --KPEVTOAASNNSNDASKYEVPKDQASKETTFT--SWEAKDFVTR-----G 144

QY 947 ALKEKLVISBLRDKDQYETP-KEKTSAVENTVSTIQSKAIDSKLURNGSINCKY 1005

QY 145 DTLVQFSKGINSKQTSQTHLVLPSHADSGQLTQVASFAFPDKKTAIRYTSRIGENQK 204

Db 1006 NTIDDLRSLKKTREBEVKRGPQKDGKQGENTVALLSLRDLRDKNGKINELN---DGR 1061

QY 205 PSRLDIDQKEITDEGBIFWVQVLTQKLTIPNGYKSTGDAF--VD---NKNLARVNLE 257

Db 1062 LNSLTIKKEPLIK---PSESKSCKHLSQDQKG-PQDPLNRIDBEDIKGRDVBLN-- 1114

QY 258 SLETTSDYARAHMSLQKVLQDNLKTYGELAFFDQIQGKLYPRHLITLKAERPKSNII 317

Db 1115 -----NYQVISE-----NKV-----TLFKNNSV 1132

QY 318 QTVERFLGSKKVKVIGEASFQDNLRNVLPIGLEKTESEATGNCDEHFTNNQVUFRTRG 377

Db 1133 TYIEMAMSHINTV-----AH3ITSN-----KNEILKSVK 1163

QY 378 ONPHOLATENTIVNPDKSLWRATPDMYTKLEEDFTYQKNSVTSQFSNKGQKVRANKNL 437

Db 1164 EDKLNIVQN-----EDYKVKVNPE--NEKQLEARGSMS- 1196

QY 438 EIPKQHNGITTEIGDNFAFRNVDQFSKTRKYLDEBILKUSTIRKIGARAFQSNMNLKSPRE 497

Db 1197 -----KLUKEV-----INKHVSMTQESTANTL---KSNAGKGE 1227

QY 498 ASEDLREKIGGAFMNMRIGTU-DLKDKLKLIGGDAFHINHIIYATLPEPSVQEOIGSAFQ 556

Db 1228 NEHDIEE-----LNKTKGQMDRIVKLUKCLAE-----LKE3VNLKDANER 1270

QY 557 NGALHLMFIGKVKITGEMAFLSNKLESWILSEQKOLKIE- 604

Db 1271 -----ANKVPEPERNTIGHVLERITVEKDAGKVVEMNSLTKIEKLIQETSD 1321

QY 605 ALSEVW--LPNPLQTIR--SEAFKEM---HLKE-VKGSSTLSQITFAFQDQNDGDKR 653

Db 1322 SQNLVLTTSITKHLENAKGVEDVIRKNEEDSIOLRERAKSLETDEM----- 1368

Db 654 FGKKVWVTRTHNSHMLANGERFIDPDQLSSTDMLRKVKGLEYSTL-----RQT 707

QY 1369 --KLUVOQWNNMQLQSAIQGNAGI-----SKELNELKQVIELLUSTNTVSSILEYVKNNSS 1420

Db 708 TQTOFREMT---TAGKALLSKENLRQEQBKQKPLQEAOFFLGRVADLQKAKAKALVTK 763

QY 1421 BSVRFSLANGERFIKAEGEKEENASARLAEEKUKQIWKDQYDSDIDKVKKE----- 1474

Db 764 KATNGHILERSIMKAVIAYNNR-----KGANVKRLEKELDILITDVE 808

Db 1475 -----GIRKEIIMKESALITFWESEKEKKQMCSSHMRNKEGKKTILKN--N 1521

QY 809 GKGKLAQATMVO---GTVLKLKPLPLRBYIGVNNVYFDK---GKLYIADMDSPTIG 859

Db 1522 GDGCKANITDSOMEVGVYNSKA---BHAFTVQAQDFTKTAFOESTIVAYVTDMDNLFN 1577

QY 860 EGOKDAGYGNPILNUEDNEGYHILAVATLADYEG----- 893

Db 1578 ESLMKEVVKVCKEKENDEAKYS---AKLKPYDGRKARVSENERKISLKEKAKVEKKE 1633

QY 894 -LYTKDILNSLODKKATRQ-----IPLAKVHRLGIFQATRNAAABDRLLPKTPKG 944

Db 1634 SSOIANDVSTKSLQDIDNCQQLDSVLNIGRVKOMALQFD---SADKSMKSVLPISEIG 1690

QY 945 YLNENPVNTY---KQKOMKQKPVYKTP1FNKALPNEKQDRAAKGHNNAE---TNNS 998

Db 1691 AEKSLDKVKAKESEYKQLETFVQ-----NEM-----SRINVERGSLTDID 1730

QY 999 VAVTPRPSQQLHKSQSDVNLPO---TSKKNPFIYELIG 1034

Db 1731 KKTIDDIENDLKKMVKQYBEGLLOKIKKENADKRIKSNFELVG 1770

RESULT 9

SCA4\_RICPR

ID SCA4\_RICPR STANDARD; PRT; 1022 AA.

AC Q92D49; Q0A336; Q9ZD48;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Antigenic heat-stable 120 kDa protein (PS120) (120 kDa antigen)  
 (Protein PS120).

DE SCA4 OR RP49; RP49.

OS Rickettsia prowazkii.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 Rickettsiaceae; Rickettsiidae; Rickettsia.

OC Rickettsiaceae; Rickettsiidae; Rickettsia.

OX NCBI\_TaxID=782;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Madrid B;

RX MEDLINE=99039499; PubMed=9823893;

RA Andersson S.G.B., Zomocodipour A., Sichler-Ponten T., Almemark U.C.M., Andersson J.O., Eriksson A.-S., Winkler H.H., Kurland C.G., Naeslund A.K., "The genome sequence of Rickettsia prowazkii and the origin of mitochondria." *Nature* 396:13-140 (1998).

RN [2]

RP SEQUENCE OF 11-1016 FROM N.A.

RA Sekerkova Z., Roux V., Raoult D.; "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the gene D, coding for an intracytoplasmic protein,"

RT 'gene D' coding for an intracytoplasmic protein;"

RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -!- CAUTION: RBP.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 234.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; AU235272; CAA14951.1; ALT FRAME.

DR EMBL; AU235272; CAA14951.1; ALT FRAME.

KW Antigen; Complete proteome.

FT CONFLICT 11 15 EFDPL -> RPLGLV (IN REF. 2).

FT CONFLICT 365 365 H -> Y (IN REF. 2).

FT CONFLICT 413 413 MISSING (IN REF. 2).

FT CONFLICT 957 957 G -> R (IN REF. 2).

SQ SEQUENCE 1022 AA; 114410 MW; 032303E3A663A9622 CRC64;

Query Match 3.8%; Score 202; DB 1; Length 1022;

Best Local Similarity 18.8%; Pred. No. 0.031; Matches 207; Conservative 156; Mismatches 400; Indels 298; Gaps 54;

Matches 207; Conservative 156; Mismatches 400; Indels 298; Gaps 54;

QY 28 LERSEBSVKOB---OTQSASDDWWFEDNEERKTVNSKENSTVETUDSDFPSDGNSNSSSK 84

QY 85 DEKQDILRKYFVNINPNEELAEQIAKEEDDRKFRAFLSNDQNYALINKAFEDTKKNIK 144

QY 85 TE-----SVWSPKQPKAKAFETVQDASNSNDASKVETVKPDTASKETLTSTWE 136

QY 145 AETVQYKVNLYSTVSYANGYQOGFOPVOWENQVSADL-RSTVVKNDGEPELCTINBITV 203

QY 137 AKDFFT--RQDTLVGFSK-----SGINKLSQTSHLVPSHADGTQLTQVASFAFPD 187

QY 204 TKDQVAKODGTQVQINSYR3INFPTKLOKANGSHNLSWVALKAQDGTPKAQKAVFT-- 261

QY 188 KKTIAIAYTSRIGEAKPSRLDIDQKELIDEGHINAYQOLTKLITPQNGKSIQGADFVN 247

QY 262 -----AHYEB--GPGK----- 272

Db 273 -----QKLEISSQPKLKEVGTGDDAVYIEH--GGIYTAV 307

QY 301 PRHLTLKLAERAFKSKRQIOTV--EFLGSKLKVIGASFODNLRLRVMPLDGLKESEAF 357

Db 308 TRGKYKEMMKEVALNKGOSVALSQTIAETDILTHVQGSPHETH--KPIIPI--QELBESI- 362

Db 358 TGNPGDEHYNNQV-----VLRRTGQPHOLATE-----NTYVNPDKSWRATP 401

Db 363 -----EQHTSQOQVPRITPENKSLQPKISQ-THOLOPOQAOQSSGTPNPVLMANALSTSMQ 416

QY 402 DM--DVTKWLBEDFTQKNSTGFSKNGKQVKURRNGLE--PKQNGITTEIGNAFRN 458

Db 417 DLINNTINSYLTQNDINKOS--DLIKEATAILNPKNSDRAEKN--TIDLAQNFNS 471

QY 459 VDQSKTRKDYLETIKLPSITRKGAFAFQSNNLKSFLEASEDLEIKEGAFMNDRIG-- 516

Db 472 KDIADA-----KVNVTNL-----LTIQNDQNTDIIKSKILEMWTAIT 512

QY 517 -----TLLKDK--LIKGDAAFH1-NHIIYAVLPLSPSVOEGRSAPRONGALHLMPIGNK 568

Db 513 INSENTELKOKOQOQILEKVKVNDIGLISKODISRVVAVD5IMD--TVIKSN-----IANE 562

QY 563 DK-----EKIFITVFDQINSYEFNSVAKOQLLSILKKAETQPLSPE--QQQLMQNDNT 617

QY 619 RERBAFKRNHLKEVKG-----SSTLSQITFNAFDQMDGDKRGKVVTRHNSHMLADG 672

Db 618 TTEHTKDTIEKVNINLPLEPLNTALKIT-----NTQYMTEN----- 654

QY 673 ERTIDPDKLSSWMDLAKVLTIEGJDYSTARQTOQFREMTTACKALIJKSNLRQGE 732

Db 655 -----VLD-----SPVQIEMSKLQIQQVTT-KTVAE SALVEPKDTKEIVKG--IKTIVTHSD 703

QY 733 KQKFLQRAOFFLGRVDDAKIAKAKEALKVTKKATG---HILLERSINKAVIAYNSAIIK 789

QY 704 TSLPLHDKVIMVOSSV--AKGIVESKNIDDLRELIATIAGLVDGIEYAKEDNAVVAHSSMIA 761







RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B.,  
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RT proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [71] SEQUENCE OF 191-195; 395-405; 545-554; 747-761 AND 811-823.

RX MEDLINE=21969647; PubMed=11973345;

RX Tran H., Pankov R., Tran S.D., Hampton B., Burgess W.H., Yamada K.M.;  
 RT "Integrin clustering induces kinase accumulation.";  
 J. Cell Sci. 115:2031-2040 (2002).

RN [8] CHROMOSOMAL LOCATION.

RX MEDLINE=96163023; PubMed=8575822;

RX Print C.G., Morris C.M., Spurr N.K., Rooke L., Krissansen G.W.;  
 RT "The CG-1 gene, a member of the kinecin and ES/130 family, maps to  
 RL human chromosome band 14q22.";

RT Immunogenetics 43:227-229(1995).

RX -1- FUNCTION: Receptor for kinecin thus involved in kinecin-driven  
 CC vesicle motility. Accumulates in integrin-based adhesion complexes  
 (IAC) upon integrin aggregation by fibronectin.

CC -1- SUBUNIT: Parallel homodimers formed between the membrane-bound and  
 CC the cytosolic form, and also between 2 cytosolic forms (By  
 CC similarity).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Vesicle membrane  
 CC protein anchored to the endoplasmic reticulum.

CC -1- ALTERNATIVE PRODUCTS:  
 Event=Alternative Splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;

CC Name=1; IsoId=Q86UP2-1; Sequence-Displayed:

CC Name=2;

CC IsoId=Q86UP2-2; Sequence=VSP 007981; VSP 007982;

CC -1- TISSUE SPECIFICITY: High levels in peripheral blood lymphocytes,  
 CC testis and ovary; lower levels in spleen, thymus, prostate, small  
 CC intestine and colon.

CC -1- SIMILARITY: Belongs to the kinecin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

DR EMBL; 222551; CRA80271.1; --.

DR EMBL; 125616; AAB65853.1; --.

DR EMBL; AY264265; AAP20418.1; --.

DR EMBL; D13629; BAA02794.1; --.

DR EMBL; AL138499; --; NOT ANNOTATED CDS.

DR EMBL; BC050555; AAH50555.1; ALT\_TERM.

DR PIR; S32763; S32763.

DR PIR; I33799; I33799.

DR Genew; HGNC:6467; KTN1.

DR MIM; 610381; --.

DR GO; GO:0005789; C: endoplasmic reticulum membrane; TAS.

DR GO; GO:0005887; C: integral to plasma membrane; TAS.

DR GO; GO:0005624; C: membrane fraction; TAS.

DR GO; GO:0006599; P: nonselective vesicle transport; TAS.

DR GO; GO:0007018; P: microtubule-based movement; ISS.

DR InterPro; IPR002117; Spectrin.

KW Endoplasmic reticulum; Signal-anchor; Transmembrane; Coiled coil;  
 Alternative splicing; Polymorphism.

FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL) SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)  
 FT TRANSMEM 7 29 (POTENTIAL).

FT DOMAIN 30 1357 ENDOPLASMIC RETICULUM (POTENTIAL).

FT DOMAIN 330 1356 COILED COIL (POTENTIAL).

FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 772 772 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 904 904 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1088 1088 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1263 1263 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1329 1329 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 1031 1059 Missing (in isoform 2). /FTid=VSP 007981.

FT VARSPLIC 1232 1259 Missing (in isoform 2). /FTid=VSP 007982.

FT VARIANT 282 282 V->M (in dsbNP:2274073). /FTid=VAR\_016206.

FT CONFLICT 15 15 S->P (IN REF. 4).

FT CONFLICT 210 210 MISSING (IN REF. 1). I->M (IN REF. 1).

FT CONFLICT 373 373 E->G (IN REF. 1).

FT CONFLICT 939 939 971FCDF8A8FCB8E CRC64;

FT SEQUENCE 1357 AA; 156274 MW; 971FCDF8A8FCB8E CRC64;

Query Match 3 / 7%; Score 199; DB 1; Length 1357;

Matched 218; Conservative 210; Mismatches 404; Indels 302; Gaps 59;

QY 2 TKKHKLQLALTTVS----VVTTSOEVG-LERBEESVKOEQTOSASEDW----FEE 50

Db 351 TKDRQQLQTMEMMTEKERSNVVITENKDRIGTLEKHNVYQETQOMQMKFQO 410

QY 51 DNERK--INVKSKNSTDUTVSDISDONSNNSSKTESWVSPPKQVKAKPPTVQEA 106

Db 411 VREQMEATAHILKQENGILRDAVSN----TNOLESKESAELNKLQRQ--DYARLUNEL 462

QY 107 SNSNSDASKWVVKPDQKASKETLSTWEAKDFVTRGDLVGFSGKSGINLQSOTSHLV 166

Db 463 TEKTKLQLQEVQVKNAEQATOLQVQAE----- 494

QY 167 PSHADGTTQITQVQASFAFTPDKTKTAEYTSRIGNGKPSRLDIDQKEITDEGETFNAQY 226

Db 495 ----RRWEEQSYI--RKRTAHE-----AAQDQSKFVAKENEWQSLH- 533

QY 227 LTKLKLTPNGYKISGQDADFVNKNIE-----VNLPESETRTISDYAFAMLSKQ 274

Db 534 -SKLT-----DTLVSQKQLEQRLMQLMSEBQVRVKNBESLQ-----MQVD 573

QY 275 VKLUDPMILKVLGELAFLFDNQIGKLU--YURHLILQLAERAFKSNII-QTEFLASLSKLV 330

Db 574 I-LEOMEALKAQIOPHSQIAQTSASVLEBHHVYI--ABKDKOIKQTEDSLASBRDL 630

QY 331 --GEASFQDNMLRNWMLPDGLEKIEKNSAFTGPNPGDEHYNNOVLRTRGTQNPHQI--ATE 386

Db 631 TSKEEELKQDQMNFLIKARVQKQLQAL-----NEQAA--AHEAEKMQ 673

QY 387 NTYVNFDPDKSLWRAFPDMDYKWLRLPEFTYQ-KNSTVGF----SNGKLOQYTRRNKLB 440

Db 674 SVTKDQK-----IRLLEQLOQHESNKMEEFKLNDQNGALKSVEQVKQLOTLV 722

QY 441 KOHNGITITELGDNARFNVUDQSKTURKYDLE--IKLPSITRKIGAFAPFOSNNLKSFEA 498

Db 723 EOPNPKDVEQM-EKCIQEKDEKDKTVEL-LFTGLIQLVATREBEEKAIATIENSL---- 775

QY 499 SEDLIEEKEGAFMNKRIGTDLKDKLKI--GDAFAFHINHYIAVLPESVQEIGRSAF 555

Db 776 TKEVQDLK--AKQNDQVSFLVSELEKKVILHEKGKLI-----KSYEELRA-- 819

QY 556 ONGALHMFIGKVKITI--GEMAFPSNKLPSVNSBQGQK-IEVOAISDMALSEBV 610

Db 820 ---ELVKVANKKEKTVQDLKOEKALKEEIGNVQLEKAQQSITISKVQELQ----- 866

QY 611 LPPNLOITREBAFKRNLKEV---KGSSTLQQTENAFDONGDGRFGKVKVWRTHNN 665

Db 867 ---NLLKREK--QMTKAVLKEERKDLANGTMQDQGNEENSIKAVQEV--AQN 918

QY 666 SMLMLDGERPFTIDPKLSSTMWDLKYLKILGL---DYSTLRTOTOTOFREMTAG 719

Db 919 LKEBASSASOF---EELELVIKEKENEELKRLFAMKERESDLSSKTLQ-----DVODEN 970

QY 720 KALLSKSNLROSEBKQFLQEAQFLGRVQDVKDIAKAKAL-----VTKAT 766





Db	QY	936 KEHTISKELVEYKERRQSIDNUVAKITERKLSSLANNYKOMAENESLIKAVEESKNESIQ 995
Db	QY	254 -NLPESLETIS-----DVAFAHMLKQVLPDNLKVIGELAFFFNDQIGK----- 297
Db	QY	1051 EYESQISLUKEKETATTANDENWNUKUSELTKTRELAELAYK--NLNU-ELEYKTLET 1107
Db	QY	996 LSMQNKIDMSOBNENQIERSIEKIEQLK---KTISDLEQRETELIKSSKID 1050
Db	QY	298 ---LYPRHLIKLAKERPKNSRITVEPLGSKLKVIGEASFQDNLRNVMPLPGLK 351
Db	QY	1051 EYESQISLUKEKETATTANDENWNUKUSELTKTRELAELAYK--NLNU-ELEYKTLET 1107
Db	QY	352 TESBAFTNGPDEHYYNNQVLRTRTGONPHOLATENTVNPDKSLIMATPDMDYTMKLE 411
Db	QY	1108 SEKALKEVGENEELKEEKI-----OLEKATEKQOLNSGR--NLESKEHE 1155
Db	QY	412 DFTYQ--KNSVTGFSNKGQLQVKVRNKNLIPKOINGITITRGDNAPR---NVDPSKTL 466
Db	QY	1156 DLAQQLKKVEEQTAKN----EZOQNELTSQUNDEITSTQONENESTICKKNDLEGEYKAM 1210
Db	QY	467 RKTDLDEEIKLPSTIRKIGAPAFOSNLKSPEASED---LEEKGAFMANRIGTLDKDK 523
Db	QY	1211 KSTPSEEDQNIKKS--EDALNQIQLKELKKNTEASLESIK-----SVESE 1256
Db	QY	524 LIKGDAAPHINITYATVLPESVOETGSAFRONGALHIMFTGNVK--TIGEMATLSNK 581
Db	QY	1257 TVKIKELQDECNTKEKEY-SELEDKLIKASEDKNISKYELQESKEKELDAKTSUQIQ 1315
Db	QY	582 LBSV--NUSBOKOIKTIEVOAFS-----DNALESEVUPLPNUQTIEBAPKNHLEKVG 633
Db	QY	1316 LEKITNLSRAKERSEBSERLSRLLKTSSEBIRKNAEQLERKUNIQINQNAFSEKERKULNQ 1375
Db	QY	634 SSTMSQITPNAFDQNDGDKRFGKKV-----WRTHHNMSLADGERFIIDPKLSSTMV 687
Db	QY	1376 SSTITQ-----EYBESKINTLEDELLRQNELEKA-----KEINTNRS 1413
Db	QY	688 DLERKVLKLTIEGLDYSTURQTTQFQR-----EMTAGKALIS--KSNLROEQK 734
Db	QY	1414 ELSKV-----SLSNDELEKONTIKSLQDEELSYKOKITRANEKULSIERDNKROLES 1468
Db	QY	735 K-ELQEAOFFGLRGRVUDLKAIAKAKEL--VTKKATKONGHLLRS-----INKA 779
Db	QY	1469 KBOQRAAQ-----ESKAYKEEGLKKEBESSKEMAKLRSKEMMKCESTISNET 1519
Db	QY	780 VLYAYNNSAIKKA-----NVGRLEKE-LDLITDLYEGKFLAQATMVQSYLL 825
Db	QY	1520 ELKSMETIRKSDEKLBSQSKSAEEDTKNLOHRSKDLISRINSEKODIE-----L 1570
Db	QY	826 KTLPLPPEVYIGANVYFDKGSKLJYALMDSTIGEOKDAYANPILANVEDNEGYHTLAV 885
Db	QY	1571 KSKURIEA-----KSGSEL-----ETVQKELANNAQKIRINA-----NWLK 1609
Db	QY	886 ATIADYE-----GLYKTDIANSLSOKKATRQIPLAKYHRLGJFQAIRNAAE 933
Db	QY	1610 SKUBDIERBLKDKQAEIKNSQBEREKLUTLSRKELE-----QE 1646
Db	QY	934 ADRLIPKPKGYLNENPVYRKCOMKEMEKNLKPVDYKTPENKALPNEKDGRAKGHNINA 993
Db	QY	1647 LDSTOQKQAKQSEBERRAEYRKEVYKS--OLDEKAMLLETQY-NDLUNKEPAWK----- 1697
Db	QY	994 ETNNSVAVTPIRSBQQLIKSQSDV-NPQTSKNNFIVEILGVSCLLFLYTAGKGR 1052
Db	QY	1698 RDEDTVKITTDOSRQEIEKLAKEDNLKAENGSLKEANEDRSEIDIMLIVYDLSKNAK 1757
Db	QY	1053 AR 1054
Db	QY	1758 YR 1759



us-10-091-007a-24.rpr

RESULT 2

T28317 ORF MSV156 hypothetical protein - *Melanoplus sanguinipes* entomopoxvirus

C;Species: *Melanoplus sanguinipes* entomopoxvirus

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000

C;Accession: T28317

R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kurish, G.F.; Rock, D.L.

J;Virrol, T.; Tulman, E.R.; Lu, Z.; Oma, E.; Kurish, G.F.; Rock, D.L.

A;Title: The genome of *Melanoplus sanguinipes* entomopoxvirus.

A;Reference number: Z20484; MURID:99102612; PMID:9847359

A;Accession: T28317

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1127 <AFO>

A;Cross-references: EMBL:AF063866; NID:94049647; PID:94049717

C;Genetics:

A;Note: MSV156

Query Match: 4-4%; Score 234; DB 2; Length 1127;

Best Local Similarity 18.2%; Pred. No. 0.003%; Indels 416; Gaps 56;

Matches 211; Conservative 189; Mismatches 346; Indels 416; Gaps 56;

QY 19 VIVSQE-VVGLERESTKQEQTQSASEDDWEEDWERKTVNS--KENSTV--DETVS 70

Db 108 ITYKKKVVYDDEEKEDKELVINTINQKNAVKINDKNNNTIHSNETITGKETLI 167

QY 71 DIFSDGNSSNSSKTESVVSDDPKQVKPPTQEAASNNSNDAKVEVPKQDTASKETL 130

Db 168 DIL-----NKLKLUVSDDKQLIB--QIVKNTNKELEPKNIDNQKEINKQDEL 216

QY 131 ETSTWEA-KDFVVRGDTLVGFSKGINKLTSQFSLWVLSHAADGTOITQVASFAFTDKK 189

Db 217 NKLDESKKEFIIKKQEE-----NK-----TIDKK 241

QY 190 TAAEYTSRIGENGKPSPLDIOKE-IDDE-----GEIENAYVQLTQLTIPNGY 236

Db 242 Q--BELIKKL--NDKEINFNIDEEQKQKLQDQINQSKINTNENITKGMMVNTETRNKISNLQ 297

QY 237 KST-GQDAFVD-----NKNIAEV-----NLPESE-TISDIAFARM 270

Db 298 NEFLINKDTIKSLSDEKQKLDELDKNTNNTISLYNKSNTKNTIQOLLESLTDFNNNT 357

QY 271 SLKQVKLPDNLKVIGELAFFFQDNGQGKXKYLPRHLIKLAAERAFKSNSNRLQTVFEGSKKVI 330

Db 358 NINELK-----SKILKLFNDNQ-----KLNDITQNN-KITDFNNSTRF 398

QY 331 GEA-----SFDQNDLNRLVMPDGKIESEA-FPGNCPGDEHYNNQVQVTRTGCNP 380

Db 399 KEKUDTEYKKIDDITKNNNQK--DEESTKKIDQTEYTKNKINKEYDIEK--NNNL 453

RESULT 3

557976 nuclear migration protein NMP1 - yeast (*Saccharomyces cerevisiae*)

N;Alternative names: protein YDR358-06; protein YDR150W

C;Species: *Saccharomyces cerevisiae*

C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 21-Jul-2000

C;Accession: S57976; S19052; S17018

R;Murphy, L.; Richards, C.; Harris, D.

A;Residues: 1-2148 <YUR>

A;Cross-references: EMBL:Z50046; NID:9899393; PID:9899399; MIPS:YDR150W

A;Experimental source: strain AB972

R;Kormaneck, J.; Schafft-Gierschlaeger, I.; Zimmerman, F.K.; Porecko, D.; Kuentzel, Mol. Gen. 20, 277-287, 1991

A;Title: Nuclear migration in *Saccharomyces cerevisiae* is controlled by the highly re

A;Reference number: S19052; MURID:92079907; PMID:1745235

A;Accession: S19052

A;Molecule type: DNA  
 A;Residues: 'I'-1569, 'V'-1571-1821, 'K'-1823-1959, 'RHL'-1963-1970, 'RN'-1973-2048, 'N'-2050-2  
 A;Cross-references: EMBL:X61236; NID:94071; PID:94072  
 A;Gene: SGD:NUM1  
 A;Map position: 4R

Query Match 4.2%; Score 224; DB 2; Length 2748;  
 Best local Similarity 20.0%; Pred. No. 0.035; Indels 258; Gaps 50;  
 Matches 228; Conservative 191; Mismatches 464; Gaps 50;

QY 7 KTLAALATTIVSVVTSQEVGLERERESVKOBQTOQASEDDWFEEDENERKINVKENSTD 66  
 Db 193 ESLKLUKTDLIELTILTKN--GMENDLS---SQRHYDKEISELKERILDLANENDRL 245  
 QY 67 ETVSDLIFSDGNNSSSKTESVVSDFPKQVKPAPVQTQ-----EASNNSDASKE 117  
 Db 246 ISVSDPTEINS-LSQNRTER-IKLQKOLDDAKASISLSSKRVQKKYQKQHTSDTIVTS 303  
 QY 118 VPVKQDPAKSKETLKRSTWKAEDFVGRGDTLVGFSKGINKLQSOTSHLVPSPHAADGQT 177  
 Db 304 DPDSEGTTSBESBDFDIDVIBDHMIEGTGSPVSEVEDISDLVKKYSEKKNMILSNDSYKLNLLQ 363  
 QY 178 QVASAPFTPKKTTAAYTSRSLIGENQKPSRSLDQKEIDGEFENAYQTKLITPGWK 237  
 Db 364 KSEASKPKDDELMTRKEVAENLNLAPNNDYNSKCEPESHLI-----KYLEASGYK 416  
 QY 238 SIGQDAFVNKNIAEVN----LPESTLEI----- 262  
 Db 417 VLPLEER-ENLNEISLNPSVTLKQKJLQKQKIPDQSTNLKEPTIPFLPILTSKID 475  
 QY 263 -----SDY-----AFAHMSLQKV-----LPDNL-----KVIGELAFFD-N 292  
 Db 476 LIITPKDYNLDFESVKNPBISBQMKKCLBANKDLSQNSICKNLEERINGCKWNLNDLYFSMW 535  
 QY 293 QI--GQKLYLPRHL---IKAERGAFKSNRQTFPLGSKLKVIGEASQFDNNL--RNV 343  
 Db 536 KIETPSKQVYLSKDAKELYDQVLDITKALEGLKNPTIDFLREK-----ASADYLLIKKD 589  
 QY 344 MLPDGLEKKESEAFTGNP---GDEHNNQV-----VLRTRGQPHOLATENTYV 391  
 Db 590 YVPSPLJYLVHAKATNHHLSDSAVEDLYKOKENDMPLELKESAKLGHTEV 649  
 QY 392 PDKSLMTRATPDMYTKWLEEDFTYQKNSVTGFSNKQKQYRNRNQLETPKQHNGITIEI 451  
 Db 650 LEKKL--EQSLEYL--VHAKATNHHLSDSAVEDLYKCKENPDMFELKEKS---AKL 701  
 QY 452 GDNNAFRVNDPFSKTLKQY-DLE-EIKLPUSTRKIGAFAFGQNNLQSFESABDLEIKEGA 509  
 Db 702 GHTVVSN-EAYSELORKXSELEKEVEQPSLAVYLVHAKATDHHLSDSAVEDLYKCKENP 760  
 QY 510 FMNRRIGTLQDKKLIGDAFHINHIIYATVLPSVEOIGRSAFRONGALHLMFIGNKV 569  
 Db 761 DVE-----FLKEKSXAG-----HTVSSEBSEYSLQR----- 787  
 QY 570 KTGEMALFLSNKLEVNLSSE-OKQLKTEIYQAFSDNALLSEVVL--PPNQTIREFAFKR 625  
 Db 788 ---KVSLEKEVEQPSLAVYLVHAKATDHHLSDSAVEDLYKCKENPDMFELKEKS 843  
 QY 626 NHILKVEVKGSSSTLQTFNAPFQDNGDKRFGKVKVWTHNISHMLADGERPFTIDPDKLST 685  
 Db 844 GH-----TVSNEAISLELKLEQPSLAVYLVHAKATDHHLSDSAVEDLYKCKENP 896  
 QY 686 MYDL--HKVLK---ILEGDYSTL-ROTTOPFREMTTAGKA---LLSSES-----NLR 729  
 Db 897 VFLKERSAKLGHTVSNEAISELEKKEQPSLAVYLVHAKATDHHLSDSAVEDLYKCK 956  
 QY 730 QCEKQRLQEAOPFLGR-VDLDKAIKAKKA-----LYTKKATKNGHL----- 772  
 Db 957 ENPDMFLKERSAKLGHTVSNEAISELEKKEQPSLAVYLVHAKATNHHLSDSAVEDL 1016  
 QY 773 -----ERSTINKAVLAVNNSAIKKANVKRLEKELDUDLVEGKGP--LAQAT 817

Db 1017 VKCKENPDMFELKERSAKLGHTVSNEAISELEKKEQPSLAVYLVHAKATNHHLSDSA 1076  
 QY 818 MVQSVYLLKTPPLPLEYTIGLNYIF--DKSGKLITYALDMSTDTGEGOKDAGYPLN--V 873  
 Db 1077 YEEVKCKENP-----DVEFLKQSAKIGHTVSVNEAISELEKKEQPSLAVY 1125  
 QY 874 DEDNGEYHTLAVATIAD--YEGIYKTDIANSIDRKAIRQIPLAKYRGLGFOAIRNA 931  
 Db 1126 EHAKVNHHL---LSDSAVEEL-VKCKENPDEFLK-----EKSAKLGHTVSNEA 1173  
 QY 932 AEADRLPKPKGVNEVNYRKQKMBEKNLKPVDYKTPFLNKALPNEKTDRAK-GHN 990  
 Db 1174 SELRKLEOPSLAYTVE--HAKATDHHLSDSAVEDLYKCKENPDEFLKEKSAK 1230  
 QY 991 I 991  
 Db 1231 V 1231

RESULT 4

B42771 reticulocyte-binding protein 2 - Plasmodium vivax (fragment)  
 C;Species: Plasmodium vivax  
 C;Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Sep-1997  
 C;Accession: B42771  
 C;Galinski, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.  
 Cell 69, 1213-1225, 1992  
 A;Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.  
 A;Reference number: A42771; NID:92315338; PMID:1617731  
 A;Accession: B42771  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-152 <GAL>  
 A;Cross-references: GB:MB80938; NID:9160627; PID:9160628  
 A;Experimental source: strain Belem, merozoites  
 C;Genetics:  
 A;Gene: KBP2

Query Match 4.2%; Score 223.5; DB 2; Length 1252;  
 Best Local Similarity 18.4%; Pred. No. 0.011; Indels 414; Gaps 52; Gaps 52;  
 Matches 219; Conservative 190; Mismatches 414; Gaps 369;

QY 11 LALTTIVSVVTSQEVGLERERESVKOBQTO-----ASEDDWFEEDENERKINVKENSTV- 65  
 Db 126 IATTKRGTSDTSQDINEL---ESTEGRVHNLQVQKQESNSMEE--MRKOILSMDLIL 180  
 QY 66 --DETVSDLIFSDG-----NSNSSSKTESV-----DPKQVKAKPPTQEASN 108  
 Db 181 NNSETIETAKESNNNTQALGPRENAKKLUKTDELLQVAMIERAKAHKONIDALEAQ 240  
 QY 109 SSNDASKVETPKDASKETLSTTWEAKDFVGRGDTLVGFSKGINKLQSOTSHLVP 168  
 Db 241 IDTVESKIEQINREMNKDEIKSISLSEMEKVKCITLISNSRGKQKIEFL----- 294  
 QY 169 HAADCQLTQVASFATPDK-----KTAIAEYTSRSLIGENQKPSRUDI--QKEIIDEGET 221  
 Db 295 -----KPKPNBESNSNKNVNINENINENI-RNSEQYQKIDDEAKQSTKVEL 339  
 QY 222 FNAYQTKLITPGWKSIGQDAFVNKNIAEVNLESLIEDVAFAHMSLQKQFLPDNL 281  
 Db 340 FHKH-----TTISNIKE-BSITLVGVTKSQKINKAEDIMKEIERNSELQTOVKQFQNL 395  
 QY 282 KVIGELAFFDNOIGKQKYLPRHLKLAERAFKSNLQITVFLGSKLKVIGEASFD--NN 339  
 Db 396 NKLNLNPHNYDN-----AEDLANDIK-----STNAKULJETNLSEVKH 433  
 QY 340 LRNVW-LPDGLEKKESEAFTGNCPGDGHYNNQVWLTGQPHOLATENTYVNPKSLWR 398  
 Db 434 LSETRNIKQGGEKISKA-----KDMQKKA-----TSENT--AEKILEK 472  
 QY 399 ATPDM-DYTKLEEPF-----YQKNSVTGFSN-----KGLQKTRNQLEIPQHNGI 446

473 VKDQDSNTVNLQQTTERNLIVTEKONLNGIDSTLTINIEGALKESKG-NYEL---GF 526  
 Db QY 3 KRHLKHTLALTIVTWSVTSQEVGLERESVKOBOTQSASEDDWEEDNER KTRVSKEN 6 22  
 447 --TTEIGDNAFRNDVDFSKTRKXDKLBKJLPSKTRKIGAFQFQNN--LKSFEASEDL 502  
 C;Species: Entamoeba histolytica  
 C;Accession: T18296  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 08-Sep-2000  
 C;Residues: 1-2139 <GUI>  
 R;Guillen, N.  
 Submitted to the EMBL Data Library, February 1997  
 A;Reference number: Z18865  
 A;Accession: T18296  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-2139  
 A;Cross-references: EMBL:1.03534; NID:91850912; PID:91850913; PIDN:AAH48065.1  
 C;Genetics: C;Gene: mncA  
 C;Superfamily: myosin heavy chain - myosin motor domain homology <MMO>  
 F;91-780/domain: myosin motor domain homology <MMO>  
 Query Match 4.0%; Score 216; DB 2; Length 2139;  
 Best Local Similarity 20.2%; Pred. No. 0.05%; Matches 260; Conservative 188; Mismatches 342; Gaps 58;

QY 447 --TTEIGDNAFRNDVDFSKTRKXDKLBKJLPSKTRKIGAFQFQNN--LKSFEASEDL 502  
 Db QY 527 LEKUEIGKRKLKVDIKKSINS-----TVGNFSSLNFPDLQYDNEKNI 573  
 Db QY 503 E--EIKEGAFMNWRIGTUDKDKLICGDAFHINHTIAIVLPESVQBIGRAFRONGAL 560  
 Db QY 574 NDYENKMGYIYNPEFGSN-----KISENIRNA-----601  
 Db QY 561 HLMFIGNKVKTIGEMAFLSNKLASVNLSHQKQKTEIYQAFSDNALSEVLPNQTLRE 620  
 Db QY 602 -----SENTSDYNSAMTIRLEAQKRV-----NLNKEE 630  
 Db QY 621 EAFKHNHLKEVKGSSTSLSQTENAFDQNDGDKFGKK--VVVTHNNHMLADGERPII 677  
 Db QY 631 BANK--YLRDVVKVYESF RPIFNMKESLDKINEMIKKQQLTNGHGNVKQVNTRELV 687  
 Db QY 678 DPDKLSSTMDLEKVKLTKTFLGDLSTLRTQTOFREMTAGK-----ANIS 724  
 Db QY 688 DENNNSDLIKQATGKNEEOKITHSTLNKAKTILGHUTSAKVGKIKTPELALTEIG 747  
 Db QY 725 KSNURQGEKOKFLQFQFL-----GRVUDKATAKAERALL-----VTKAT 766  
 Db QY 748 DAKLTKTQAEFLFESKRNVVLTEENMSKNTNEDLVHNIDAYKVAELIAHSPEIDTQK 807  
 Db QY 767 KNGHJLERSN--KAFL--AYNS-----AIIKANVYKLEVELDLTDVEGK 810  
 Db QY 808 DSSKLUJEMGNOIYLVKMLNQYKOKKISSIKSKKEAVSKVIGNSKHSLSKLT-----861  
 Db QY 811 GPLAQATMVQVGLKTPPLPEVYIGLNVFDESGKLYALMSDTGEGQDAKGPI 870  
 Db QY 862 -----CSDKSYDNIALEKOTEL\_QNLTNSFTQBK 890  
 Db QY 871 LNVDEN-----EGYTLLAVTADLADYEGLYTK-----897  
 Db QY 891 TNTNSPSKLERKIKTPESKRALKLKTIEGEMNALKASSDNHEHVOSSERVPNLSIEKE 950  
 Db QY 898 ---DTINSSDK-KKAIRQPLAKYHRUGIFQARNA-----AEDRLJPKTPKG 945  
 Db QY 951 ETDIDPSNLNTALDELKKLKGRTCEVSKY-----KLKDVTKEISDDTELINTENKVAY 1004  
 Db QY 946 LNEVNPYRKOMENAKPVDYKTPEN-KALPN\_EKUDGJRAAKHN-INAETNSVAV 1001  
 Db QY 1005 L---AYIKKNEYDFTQDVLTLNFTENKQVSNEHPTNFDSNKSSEELTKAVTDSKII 1060  
 Db QY 1002 TPIRSBQOLHKSQSDVNLPTQSSK-----NFTIELGLYSL 1038  
 Db QY 1061 SKLKGVILIEVENTENTIETSSAKKEAELNNEKKTSNLVYQTSNEKTL 1112  
 Db QY 1118 LTKTKDADLKKLKGRTCEVSKY-----KLKDVTKEISDDTELINTENKVAY 1177  
 Db QY 119 IAE-----NGKSRSLDQKEID-----EGIFNAY-----DUTKLITIPNGY 236  
 Db QY 1238 EKMKKALQEEKVENVESSKNSTIDKKKEDNUKDQKCLDDMT-----ADNE-----1284  
 Db QY 306 KLAERAFKSNRLOTVERFLGSKKRVVIGEASF-----ODNNRNVMF-DGJKEKSHF 357  
 Db QY 1285 KLIKAKA-KDLEAQLENEVQDNEKAVADAEULLKKKQASDKEINSLKLEALETAKSVUE 1343  
 Db QY 358 TGPQGDEHTNNVWLRTRGTPQPHOLATENTVNPDKSLIWRATPDM-----404  
 Db QY 1344 SKRKDSE--NEKAALSEEDIDORANEKL-----NIQADRKTADLOBANEKKAEEQR 1395  
 Db QY 405 -----YTKWLE-----BDFTY-----OKNSVYCFSNKG-----QKVRNKN 436  
 Db QY 1396 KDLVADNKONTKUTLEETKARDBENTYKVNENYKVLKKEADLBEANNDIDKDRMK 1455  
 Db QY 437 LEIPKQHNGITTEIGDNAF-----RNVDQSKTURKYDIEE-----KLPISTIR 481  
 Db QY 1456 KQVKKLEGKIKETDKUNAKLAKDSIFTAKQKOSDADLEELNKTEHDEVAKLNQIT 1515  
 Db QY 482 KI---GAFAPQSNLNLKSFEEADELEKE-----GAFMNRNRTGTLIDKDKJIKIGAAF 532  
 Db QY 1516 KLTRDQSAEEBELNLRs-KADDKKKISELLEQVNELESRPITGWNADENFIKIRDA- 1572  
 Db QY 533 HINHYIATWLPESYQELGRSAFFONGAHIMFLGKVK-TIGEMAFISN-----580  
 Db QY 1573 QIADINKALEMKVYVNNQIQLATQNTKELKAKNDLTSKIEDE-----1632  
 Db QY 581 -----KLESVNLSEQKQKLTIEVQ---PSDNALSEV-----LPPNQQTIREAF 623  
 Db QY 1633 ADKAVSEQTIKRKQKLEEEVKLTTEIOLAKFOINAPSSVQARERKORLESIDALKEOLB 1692  
 Db QY 624 KRN-----HILKEVKGSTSLSQTENAFDQNDGDKFGKVKVVTTHNNH-----668  
 Db QY 1693 QERTTAANAAERKKIQAELDEVKENLEDTVNGRE--KLVAKNSENDAEELSIKEKKA 1749  
 Db QY 669 LADG-ERFTIDPDKLSSTMDLEKVKLTKTFLGDLSTLRTQTOFREMTAGKLL--SK 725  
 Db QY 1750 LDEDETEKITDDNNKUSEEIDSLDRKNA-----LDSKDSVMSKKEKFQDELKVTDALETEK 1807  
 Db QY 726 SN---LR\_OGEKOKFLQFQOFFLGRVUDLKAIAKAKLVTGKATGKCHLRSINKAV 780  
 Db QY 1808 RKAETMRLKRLKEAEEVQVRLAQLDQKERA--KATDKRADELSKIM 1863  
 Db QY 781 LAYNNSAIKGANV-RLEKEELDUTDVLGKGLAQ--ATMVGCVYIUKTPPLPEYYIG 1837  
 Db QY 1864 NELD-----DVKDQDKRQDPLADKEDELATDQKTYLKVOKSKVFSRDSIOMEOQD 1916

QY 838 INVYFDKSSKLIYALDMSDTIGEORDAYGNPLAV--DEDNEGHITA--VATLADY 891  
Db 1917 L---EKAGR-----AKAQKQXAEKQKLOQLOQNDDEEYKPAKDKINTLSAQ 1963  
QY 892 EGLYIKDIINSSLQKIKAIROQPLAKYHUGI-FOAIRRAAE--ADRILRKPKGYIN 947  
Db 1964 ----KDDIQKELEKERGIRQDSEKEVQRLVKQBLETRKVAEVGANVIKVKAYEA 2018  
QY 948 EVPRY-----RKKQMEKNUKPVDPYKIPFINKALPKEVKDRAAKGHNIAETMS 998  
Db 2019 EIEELUTTEADALKAKMKKEKAKTSQKUDLEOKTI-----ADYETEBSAFTEIG-- 2070  
QY 999 WAVTPIRSEQOLHQSQSDNLPQT 1022  
Db 2071 -----KTOQABLKXVQQVRDDET 2083

RESULT 6

T50073 myosin-like coiled-coil protein sp8 [imported] - fission yeast (*Schizosaccharomyces pombe*)  
C;Species: *Shizosaccharomyces pombe*  
C;Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000  
C;Accession: T50073  
R;McDouall, R.C.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.  
Submitted to the EMBL Data Library, December 1999  
A;Reference number: Z25034  
A;Accession: T50073  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Cross-references: EMBI:AU133357; PIDN:CA06241.1; GSPDB:GN00066; SPDB:SPAC1486.04C  
A;Experimental source: Strain 972h(-); cosmid c1486  
A;Gene: SPDB-SPAC1486.04C  
A;Map position: 1

Query Match 4.0%; Score 215.5; DB 2; Length 1727;  
Best Local Similarity 19.3%; Pred. No. 0.043; Mismatches 231; Conservatve 198; Indels 343; Gaps 57;  
Matches 231; Conservative 198; Mismatches 422; Indels 343; Gaps 57;

QY 12 ALTTWVSVTYSQEVIGLEREESVKQEQTASED----DWFEEDNERKTNVSKENSTWD 66  
Db 116 ALTT-----ENGSIRANSELQEQSKIASEQSIANDQIRALQNEENSHLGQVQASH 167  
QY 67 ETVSD-----LFSQDGANNNSKKTESVSDPKQVKPKAKPEVQEAANSNDASKV 118  
Db 168 QALSDIBERKKQHMPA--SSSSRVEKEELVQE-----KSLVSDLQLQSDHSKV-- 215

QY 119 PKQDFTASKETLETSTWEAKDFVTRGDTLPGPSKG--INKLSTPSHVLPSHADG-- 173  
Db 216 -----CEKLEVSSRQVQDPLEK--LAGLAQCTEALNEKIQOLFEQQRKSNYSSDGNIS 264  
QY 174 -----TQLTQVASFAFTDKKTAIAEYTSRILGENGKPSRDLID--QKEIILDEGF 222  
Db 265 KILETDPSTSKELEBVERETQKRLTAWE-----SKSELLOSEVAALQEKLSQSQSLY 316  
QY 223 NAYQLTK-----LTIPIGYKSTGQ--DAFVQKNAEVN-----LPE 257  
Db 317 N--NTEELNNKQQLLISENSLSRELQKVDWSVSLQVTKENKNTSVAGVGLPSLQ 374  
QY 258 SLETIS--DYFAHMSLQKQVLPDN-----KYGELAAPPDNOIGSKYLYPLRHLKL 307  
Db 375 KLSAVVNPEPSFTKVISSDNMKLQKQVSSLQKQLDRLTNKFSSFCQV-KQRIP--VVKQ 430  
QY 308 AERAFKGSNRQTVFEGSKLKVIGEASPFQDNLRLVNUMLPQLEKESEATGNPCDEHNT 367  
Db 431 ORSEIVTRNNTY-MNFLSESLT-----SNNLNTKVOAELISTKQHQAC-----YL 475  
QY 368 NOVQVTRTRGONPHO--LATENTYQNDPSLWRATPDMDTKWLJEDFTYQKNSVTG-- 422  
Db 476 QLTASSTQCSLISREVICLMLDHNTES--RNPATVQVA-LDE--YAQPSTAGE 529

QY 423 -FSNKGLQ----KVRNKNLEIPKOHNGITITEGDNAPRNVDFOQSKLRYDLEIKL 476

RESULT 7

A95250 choline binding protein PcpA [imported] - *Streptococcus pneumoniae* (strain TIGR4)  
C;Species: *Streptococcus pneumoniae*  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C;Accession: A95250  
R;Tettelein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heaton, J.D.; Umayam, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, T.; Hickey, E.K.; Holt, T.E.; Science, 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.;Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.  
A;Reference number: A95000; NID: 21357209; PMID:11463916  
A;Accession: A95250  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: GB:AB005672; PIDN:AKT76194.1; PID:gi14973649; GSPDB:GN00164; TIGR:S  
A;Genetics:  
C;Experimental source: strain TIGR4  
A;Gene: SP2136

Query Match 4.0%; Score 212.5; DB 2; Length 621;  
Best Local Similarity 21.9%; Pred. No. 0.013; Mismatches 149; Indels 159; Gaps 20;

Matches 107; Conservative 73; Mismatches 149; Indels 159; Gaps 20;



QY 1039 CLLEFLV 1044  
 QY :::  
 Db 1059 LKTYIV 1064

RESULT 9

A90551 conserved hypothetical protein MPYU\_3130 [imported] - *Mycoplasma pulmonis* (strain UAB CT

C-Species: *Mycoplasma pulmonis*  
 C-ID: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C-Accession: A90551  
 R-Chambaud, I., Heilig, R.; Perris, S.; Barbe, V.; Samson, D.; Galisson, P.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001.  
 A-TITLE: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*  
 A-Reference number: A95512; MUID:21267165; PMID:11353084  
 A-Accession: A90551  
 A-Status: preliminary  
 A-Molecule type: DNA  
 A-Residues: 1-2819 <KUR>  
 A-Cross-references: GB:AL445566; PID:914089727; PIDN:CAC13486.1; GSPDB:GN00153  
 A-Experimental source: strain UAB CTIP  
 C;Genetics:  
 A-GENE: MPYU\_3130  
 A-Genetic code: SG3

Query Match 4.0%; Score 212.5; DB 2; Length 2819;  
 Best Local Similarity 18.2%; Pred. No. 0.12; Matches 246; Conservative 194; Mismatches 444; Indels 467; Gaps 56;

QY 1 MTKKHILKTLA---LTVSVWVYQSEYTLREESV---KQEQTQASASDDWEEED 51  
 3 MKNKTLIGLGLASWAGIITMSAIIPLK-TGKGDVAVVNDKNKDSSETSGSYSSDDFKDN 60

QY 52 NERKINVSKENSTDETISDLFSDGNSNNSKTESVSDPKQVPAKEVTOBASNSN 111  
 61 THSPDDDKXANQ-----NKDKKSQSRSSSD-----LENKONSEN 96

QY 112 DASKVEVPKQDASKETKETLTSTWEAKD---FVTGRGDTL-VGF-----SKS 153  
 97 NFDSGFIDMNSTLSNDKRL---IKDKIRFVFFGDSIAGFNAKMLDVPGHEDPNTK 150

QY 154 GINKLQSORTSHLVLPSHAAGTQLTQVAFPAFTPKKTAEAETSRGL---NGKSR 207  
 151 QITGLSYSSYIADYVINDLPNPKESPKNFAFS---GTTKDWYDFNNDTVPFNSTKE 207

QY 208 LDIQKEITIDEGEIFNAQYKLTIPNGKSIGQDAFVNKNIAEVNIPESLETISDYL 267  
 208 -DLSQREDP?-----LRKLT-----NSNLILLNLGAN-----DF 234

QY 268 AHMSLIKQVILPDPNQKVIGELAFFFQIGKLYLP---RHLI---KLABRAFKSNRI 317  
 235 MSLAFKEFKEKDVLISLMLMKDASINDLIG-VFLPILSSIRHEMGIYKELISHIRKWN 293

QY 318 QTVERLIGSLI-----KVIGEASRDN-----NLRNMLPQGPLEKESEATG 359  
 294 VTMILIGYPMAPKLFNMVKLIGESVKIGNDTLIGFLINFINNSIKQDVLVENVN 353

QY 360 NPGD----EYHN-----NOVVLRTTRGPNPHQATE-NTYVMPDK--SLWRAT- 400  
 354 AVDDPLWLKEKNDFTDVAIDHPEIGYKMAOBFILKLAIDSYNNNESETNSMWNY 413

Db 401 PDMYTKWLE-EDTYQKNS---VIGESNKGLQKVRERKNEBIPKQHNGITIEIGN 454  
 414 LSYDHGKRSRLEFFENKONKSEIITIVYGLNSALPSKQNDKIALYESTKNTDVYI--- 470

QY 455 AFRAVDQFQSKRLKRDLEIQLPSTIRKIGAFAGAFOSNNKSFSEADEEIEKAGFMNR 514  
 471 LFRFSDI PFKLSSKAVE-----FLINSSITKTPBEEDEI---KCFLPFANNE 513

Db 515 IGTUDIKDKLKLIGDAFAFHINHYAIVLPEPSVBIGRAFRONGALHNFIGNKVKTE 574  
 514 ANYQKLFELITKSRPVMNINRSDTFLDLDNGQAGIQLBISHLMTFRR---E 568

QY 575 MARLNSKQ-----ESVNLS-----BOKQKTEVQAFSDNALS 607  
 QY :::  
 Db 569 ISNLNDNYLGLRFESSSSPFYVDPETNLSFKD1KKKLNFLQKSELKTTLSKFTGAFIQ 628

QY 608 EVULPPNQOTIREEAKRNHLKVKGGSTSLSQTPNFMFDQNDGDKRECKKVUR---- 611  
 Db 629 NKLLSHNSEL-----VLRIVSGEKTPALL----EKLIDDIFDNKAVYQVEKIE 674

QY 662 -----TANNSHLMLADGERFIIID---PD----- 680

Db 675 DELFAPVUNNEKAITDYEPRFVNLWAKPENFKALVSGISQFLNNSNEVISTQTSQFY 734

QY 681 KLSSTMVD-----LEKVLK-----ITE-----G 698

Db 735 EVSKTLVNLVDTBELLERKVLKNFIEKSSAQIJKLUSNPKEFNLKLOKSIISTFGFTND 794

QY 699 LDYSTLRCPTQ-----QFREMTTAKALLSKSLRQGEKKEKQLEAQFFGRVLDK 751

Db 795 LGWSFPLKQVVKSKLSSVYBORDKPTSSLSKEVWSVFKSEVNVKNLISAGTYLSNTELFK 854

QY 752 AIKAKAKALVTK--KATKNGHLLERSIN-----KAVLAIN-----NSAIKKA 791

Db 855 ENDVFKGFSHTIINFCEKPELNLKIDYVINSWEISDQYDAFEEVFLKNUISWTK 914

Db 945 AGFELTNSVSESTKRNIFNTKLTIDSN-----TIDALND---TFVNSKTLIEDKKNF 995

QY 850 YALDMSPDTIGEGOKDAYNPILNVEDNEEGYHFLAVATLADYEGLYIKDILMSLDK--- 906

Db 907 ---IKATRQIPLAKYHRIGIQATRNNAAEAR----- 936

Db 996 TASLNLQIFASDDNFFSVKALISNKLKOKAFVNQYKODITNWLKELLEEEFKENL 1055

QY 937 ---LIPKTPKGY-----LNEVNRKKQOMEKNULP----- 963

Db 1056 YSPFMMPKTSISELFENNPEALKNTTILRNALKNDPFPINKMDIIFDNLKQKLENYG 1115

QY 964 -----VDKYPFPIENGALPNEK-----VGDRAKCHNINATNNNSVA 1000

Db 1116 QVGLLILKNDSDAIATKVQAIKKDVNVNANFQDVLKVILVNSQTKKKFMGPFLNADBTNKIS 1175

QY 1001 VTPIRSEPOLHSQSDVNLQPLQTSKNNFIVE 1031

Db 1176 TDITKLQFENNNFAKSLALLATKDSVID 1206

RESULT 10

T39068 coiled coil protein - fission yeast (*Schizosaccharomyces pombe*)  
 C-Species: *Schizosaccharomyces pombe*  
 C-Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C-Accession: T39068  
 R-Murphy, L., Harris, D.; Barrell, B.G.; Rajandream, M.A.; Connor, R.E.; Wood, V. submitted to the EMBL Data Library, August 1997

A-Reference number: Z21825

A-Accession: T39068

A-Status: preliminary; translated from GB/EMBL/DBJ

A-Molecule type: DNA

A-Residues: 1-1208 <KUR>

A-Cross-references: EMBL:ZB1317; PIDN:CA03608.1; GSPDB:GN00066; SPDB:SPAC6G9.06C

C;Genetics:  
 A-GENE: SPDB:SPAC6G9.06C  
 A-Map position: 1

Query Match 3.9%; Score 209; DB 2; Length 1208;  
 Best Local Similarity 19.4%; Pred. No. 0.051;  
 Matches 220; Conservative 170; Mismatches 367; Indels 378; Gaps 52;

QY	20 TYSQEVTYGLERESVKQEQQOSA-----SEDDWFEDNERKTYNSKENSTVDETVSDLTS 74	Db	1045 QSKTNNDH---LSTILERNRKEYKSL-----LDDYNOLR---ARYKNL 1082
Db	93 SFDPNSYGL---SAISKQATOEAISLTSQGNDSY---DVSKLTUDISK-NSBIDHTGEL-- 143		
QY	75 DGNSNMSSSKRESVSD-----PKQVKA---KPEVQEA 108		
Db	144 PANAALTTRREOBVKLVEKVSRENRGFLRITKTCLEKLESMAPEQIKBAVNKLVEHARAN 203		
QY	109 SSNDASKVEVPQDKTASKELETSTWEAKDFVTRGDTLVGFSKGSGINKUSQ-----T 161		
Db	204 LQLQKRTESLQSKEDPKFKLE---EKVYDLSK-----VNDYEQSQVQVFT 248		
QY	162 SHIVLPSHAADGTLQVA-SFAFTPDKTAJAY-----TSRUGENGKPSRUDICK 213		
Db	249 ERIRFLENALEKVKOREKDSLSTEMEDKSNKEVDYKYEIROLQNRLDELSR-ELDVAQ- 305		
QY	214 ELTDEBIRFENAYQVLTKLUTPGYKSTQGDAFVDNKIAEVLPE-----H- 257		
Db	306 DLITEKEB---DEIATLKRQTEKENNSAFENEENSSYVHQLQEDYALQACDDEFADRI 361		
QY	258 ---SLETISDYAFAHMSLKVOKKLDNLKIGELAFFDNOIGGKYLPRHLKLAERAF 312		
Db	362 QVLTABLEKEKENOIMHESEASIGHTSMOV-----H- 400		
QY	313 KSNRIQTEVFGLSKLKVIGE---ASFQDNLRNVLNMPDGLEKIESEAFGTGNPGDEHYNQV 370		
Db	401 KAN---SEIEFLHQDISRMNEBCKNFSIDMLQPSLSEERDVKQT---LEDDNNSL 454		
QY	371 VLRTRT-GQNPHOLATTTYNDPKSIIWRATEDMDYTKWLRHDFTYQKNSVTGFSNKGQ 429		
Db	455 RUMISSIONQTESRIONREIDEENKHLILASKNSDKALE-----TNIRLQ 502		
QY	430 KURRNKMEKIPKOHNGTITTEGDNARFN-----VDFQSTKLRYKDE 472		
Db	503 RV---TKELTBELTMKNSNDLNELHDRLREBENGLTJKTSITKEDRDLINELBQRIKSYEVN 560		
QY	473 EIKLPSTIRK-----IGAFAFOSNNIKSPSEASDURHKEGAFMANRIGT 517		
Db	561 VSRLNGTIDEYRNKLKOPKEETYNEVMNAPFOYKNDLARRFHESINKLQDREKELTSN--- 616		
QY	518 IUDKDKLTKIGDAFHINHIIYAVLPESSVQETGRSAFRQNGALHMFINGKVTGEMAF 577		
Db	617 LEKGKQVLT-----SSREVAMLEKERSSIKVIGN----- 648		
QY	578 LSKTKLSTPNLSB-----OKOLKLT---SYOAFSDNALSVLPNPLQTREAFPKRN 626		
Db	649 -AKDLDNTNLMETLNDKTSVLRQLTDKDELDVSEBEREAIVAGOKLSASPELMSENK 707		
QY	627 HLKVEKGSSSTLSQITFNAFDQNDGKR---PGKTVW---VRTHNNSHMLADERFI 676		
Db	708 QALELKYSLSKNEL-INA---QNLDRBEEBELSKKLUPEERKIRSGNDDEEKNBIN 764		
QY	677 ID---PDKLUS-----STNDVILEKU-----LIKIEGLDYLSTIR- 705		
Db	765 IASBLADILKLAQTRHLESKMEQDQLVHILNGRIBEEANIEENAVKRIKLLNGCDSVSI 824		
QY	706 -----QTTQTCRREM-----TTAGKALLSK 725		
Db	825 LQIVSQIHEFVNQOITIRSLKQELRHDFVQSGKKQOELSRSFEKFGFUGTETKHDIAQ 884		
QY	726 SNRQGEKQKELB-AOFPLGRD-----ADKATAKAELVTKATK 767		
Db	885 RNRNVESEKNDLENAQOFPSSPDRKNGYLYPSEHTSKIELEYKTIEDUKLAL---ODEIK 942		
QY	768 NGHLERSINKALVAINSAIK-KANTVGLERELDUD-----LVEGGK 811		
Db	943 NRNLIMDDIS---SYNKQTTKUQEKIKMERSILIDELSESYNSNQNPYONLVDQDN 998		
QY	812 PLAQATWQGVYVLLKTPILPEVYIGNLY---FDKSKLILVA-LDMSDTIGBQOKDAY 866		
Db	999 ELLER---LKEIQLKELVYNNFPMKQEMLTSNVTBESQMLKTLRBL 1044		
QY	867 GNPITLNVDBDNEGHTLAVATLADYEGIYKIDILASSIKIKAROPIPLAKYHL 921		

RESULT 11

myoIn-like protein MUP1 - Yeast (*Saccharomyces cerevisiae*)  
N;Alternate names: protein YKR095w; protein YKR415  
C;Species: *Saccharomyces cerevisiae*  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 21-Jul-2000  
C;Accession: S38173; S30647; S31207  
C;Baldron, V.; Ballesta, J. P.G.; Bou, G.; del Rey, P.; Esteban, P.F.; Garcia-Cantalejo, A;Reference number: S38158  
A;Accession: S38173  
A;Molecule type: DNA  
A;Residues: 1-1875 <BAL>

A;Cross-references: EMBL:Z82120; NID:9486596; PID:9486587; MIPS:YKR095w  
A;Experimental source: strain S288C  
R;Bou, G.; Esteban, P.F.; Baldron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; Yeast 9' 1349-1354 1993  
R;Koellling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.  
Mol. Gen. 237, 359-369, 1993  
A;Title: The complete sequence of a 15 820 bp segment of *Saccharomyces cerevisiae* chromosome 13  
A;Reference number: S40644; NID:94205265; PMID:8154186  
A;Accession: S406447  
A;Molecule type: DNA  
A;Residues: 1-1875 <BOU>  
A;Cross-references: EMBL:Z73541; NID:9450550; PID:9450554  
A;Experimental source: strain S288C  
R;Bou, G.; Esteban, P.F.; Baldron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; Yeast 9' 1349-1354 1993  
A;Title: A new yeast gene with a myoIn-like heptad repeat structure.  
A;Reference number: S31207; NID:93247549; PMID:8483450  
A;Molecule type: DNA  
A;Residues: 1-300; A'; 302-1875 <KOB>  
A;Cross-references: EMBL:101892; NID:97171958; PID:AAA34783-1; PID:97171959  
C;Genetics:  
A;Cross-references: SGD:MLP1  
A;Map position: 11R

Query Match

Best Local Similarity 3.9%; Score 209; DB 2; Length 1875; Matches 221; Conservative 194; Mismatches 375; Index 362; Gaps 55;

578 LSKTKLSTPNLSB-----OKOLKLT---SYOAFSDNALSVLPNPLQTREAFPKRN 626

Db 649 -AKDLDNTNLMETLNDKTSVLRQLTDKDELDVSEBEREAIVAGOKLSASPELMSENK 707

QY 627 HLKVEKGSSSTLSQITFNAFDQNDGKR---PGKTVW---VRTHNNSHMLADERFI 676

Db 708 QALELKYSLSKNEL-INA---QNLDRBEEBELSKKLUPEERKIRSGNDDEEKNBIN 764

QY 677 ID---PDKLUS-----STNDVILEKU-----LIKIEGLDYLSTIR- 705

Db 765 IASBLADILKLAQTRHLESKMEQDQLVHILNGRIBEEANIEENAVKRIKLLNGCDSVSI 824

QY 706 -----QTTQTCRREM-----TTAGKALLSK 725

Db 825 LQIVSQIHEFVNQOITIRSLKQELRHDFVQSGKKQOELSRSFEKFGFUGTETKHDIAQ 884

QY 726 SNRQGEKQKELB-AOFPLGRD-----ADKATAKAELVTKATK 767

Db 885 RNRNVESEKNDLENAQOFPSSPDRKNGYLYPSEHTSKIELEYKTIEDUKLAL---ODEIK 942

QY 768 NGHLERSINKALVAINSAIK-KANTVGLERELDUD-----LVEGGK 811

Db 943 NRNLIMDDIS---SYNKQTTKUQEKIKMERSILIDELSESYNSNQNPYONLVDQDN 998

QY 812 PLAQATWQGVYVLLKTPILPEVYIGNLY---FDKSKLILVA-LDMSDTIGBQOKDAY 866

Db 999 ELLER---LKEIQLKELVYNNFPMKQEMLTSNVTBESQMLKTLRBL 1044

QY 867 GNPITLNVDBDNEGHTLAVATLADYEGIYKIDILASSIKIKAROPIPLAKYHL 921

Db 1045 QSKTNNDH---LSTILERNRKEYKSL-----LDDYNOLR---ARYKNL 1082

Db 387 -----IFLKKQLIK---ERRTEKHLOQIETTIVF 413

Db 323 LGSKKVJIGEASFOQNNLRNVLNMPDGLEKIESEAFGTGPDHYNQ---VUVRTTQNP 380

Db 414 LEHKVPPIN- SFKERT-----DMLNELNVAALLJELHTSNEKNAVKELNAKQ 461



RESULT 13  
 uncharacterized protein, probably surface-located [imported] - Clostridium acetobutylicum  
 C;Species: Clostridium acetobutylicum  
 C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C;Accession: C97033  
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
 A;Reference number: A95900; MUID:21359325; PMID:21359325  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: <KUR>  
 A;Cross-references: GB:AE01437; PIDN:AAK79054.1; PID:g15023995; GSPDB:gn00168  
 C;Experimental source: Clostridium acetobutylicum ATCC824  
 C;Genetics: CAC1080

Query Match 3.9%; Score 208.5; DB 2; Length 1227;  
 Best Local Similarity 18.9%; Pred. No. 0.055; DB Matches 221; Conservative 164; Mismatches 363; Indels 419; Gaps 54; DB 65 N-----DSNSKVSSNENGVNSNKNSNPKVSSSETOINKNTLQVQNNKS 112

Query 62 NSTVDEFTVSDLFSDGNSNNSSKTESVSPDKQVKAKPEVTEASNS-----NDAS 114  
 Db 115 KVEVPKDTASKETELTSTBAKDFVTRGDTLVGFSGINKLSQTSHLVLPSHADST 174

Query 113 VLAASVNDDEVKK--INSGVQTSVDIAIGITKV----- 144  
 Db 175 QLTQVASFAFTPDKKVAAEYTSRSLRGENGKRSRDLDOKEID-----EGBIFNA 224

Query 145 -----KDTLDINKAIVDARSAGTDLSEWYIDI 175

Query 225 YOLKTULITPNGYKSGCDAFYDKNKNIAEVNLPESLTITSDYAFARMSLKQVLPDNLKV1 284  
 Db 176 VSQTAEAT-----QFAP-----RNIDGVANVSDT-----LL 203

Query 285 GELAFFDNDQIGGKYLPRHLIKLAAFRKFSKSNRIQTEFLGSKLKVIGEASFQDNNLRNVM 344  
 Db 204 G-ATFDNAN-----LDSVNUKYFHYKRYT-----IKFKDV-----ATKTSNALKN- 245

Query 345 LPDGLEKIESEAFTGNPGDDEHNVNQVVRTRTGQNPQHQLATENTYVNDKSLWRATPDM 404  
 Db 246 -----NN-----GQG----- 254

Query 405 YTKWLRDFTYQKNSVQGFSNKGLQKTRNKALEIPKONG--ITTEIGDN----- 454  
 Db 255 YTA-LE-----VSGVTPQYLDV--NKNIVKEKONKGRDLTITEISDAASTIARI 302

Query 455 --AFRNDFQSKLIRK-----LIEEIKL-----ESTIRKIGAFFQSN 491

Db 303 NTFALDNMDAGVATLEDYQIAIGANNVPOHADYVNSLMDQRNGDVSEADGINTIMTYIN 362

Query 492 NLKSFEEASDLERIKEGFMNNGRIGID-----LKDQLKIGDAFHINHYIAVLP 543  
 Db 363 NINSGVGETD-DYINSHAVDSNE-GNIDYDILNANTIEKKTAKQD-----LTIP 410

RESULT 14  
 surface-located membrane protein 1 (lmp1) homolog - Lyme disease spirochete  
 C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
 C;Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 11-Jan-2000  
 C;Accession: B70126  
 R;Fraser, C.M.; Cajens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, S.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugay, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Bowman, C.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugay, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Nature 390, 580-585, 1997  
 A;Authors: Smith, H.O.; Venter, J.C.  
 A;Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.  
 A;Reference number: A70100; MUID:98065943; PMID:9403885  
 A;Accession: B70126  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-119 <KLE>  
 A;Cross-references: GB:AB001131; GB:AE000783; NID:92688098; PIDN:AA66595.1; PID:9268888  
 A;Experimental source: strain B31  
 C;Superfamily: unassigned tetrastricopeptide repeat proteins; tetrastricopeptide repeat

Query Match 3.9%; Score 207.5; DB 2; Length 1119;  
 Best Local Similarity 19.4%; Pred. No. 0.053; DB Matches 223; Conservative 184; Mismatches 399; Indels 341; Gaps 59;

us-10-091-007a-24.rpr

QY 35 KOBOTOSASED-DWFEEDNERKTKVSKENSTV-----DETVSDLFDGNS----- 78  
 QY |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 94 KDLORDSAMQILDKTKGQDNTKTAVN-EPDIAFNRYKRDSTTENYSDRNDVGIRE 152  
 QY 79 -----NNSSKTSVWSPK-----QVKKPEV-QEAN 108  
 Db 153 DISEFKSKLPEKIKPNTNPKPEEDQIOPPNPKLSVNDQXNLENLEKUKKNNLGKSNEN 212  
 QY 109 SSNDAKSVKPVQDQDQASKETLSTWEAKD---FVTGDTLVGFSKGINKS---QT 161  
 Db 213 ILNQSKINDKONTNLSEKNSNMLKPDNSKSNANT-----TSIKKISSNCKE 255  
 QY 162 SHLVLPISHADG-----TOLTOVASFAFPDKKTAIAHYSRSLGKPSRLDQKEI 215  
 Db 267 SELSPPSQTIGKLYRPSYLVKELYETLDDING---RTVIGKN---RL---KEL 314  
 QY 215 IDEBRIFNAQ-----LTKLTPNGYKSGQDARFDVNKAEN- 254  
 Db 372 EDKRPQYLDLDSKSH---SXPDLNTKSRSQAKDNE- FLKN-----PND-A 418  
 QY 306 KLAERPAFKSRIQTVFPLSKLVIGEASFDQDNLRNVMULDGPLEKTESAFTGNGDHE 365  
 Db 419 QASKTLAQANKIQHEDJUSKVKHSKIPLENTKRSRQAKD---LNEFLKPNPNDQ 473  
 QY 365 YNNQVYLRTTGTQPHOLAMENTYNP----DLSIWRATPDMYTKMLEDFTYQKSY 420  
 Db 474 ASKTLAQANKI-QHEDLDSKVKHSKIPLENTKRSRQAKD---EP-LKNNN 524  
 QY 421 TGFPSNKGLOKVRKRNQLEPKQHNGITTEBIDMAFRNVDQSK- TURKYDIERIK-LP 477  
 Db 525 DAQASKTLAQANKIOLHE-----DIKSVKHSKIPLENTKRSRQ 563  
 QY 478 STIRKGAFAPQSNLKLSPASEDDEIIEKGAFMNRIGHTL-DIKDKLKLIGDAFHINH 536  
 Db 564 QAIKDLNEFL---KONPNDQASKLQA-----NKQHEDLDSK-----VHS 604  
 QY 537 IYATWL----PESVQBLGSAFRONG---ALHLMITGNKRTIGMAFLSNKLESSV 585  
 Db 605 IKPDLENTLKSROQAKD---NEFLKNNPNDQASKTLAQANKIQHEDLDSKHSIKEI 662  
 QY 586 NLSEDKQKLTIEVAFSDNALSEYVLT---PNLQTR- EFAFKENHLKVKG----- 633  
 Db 663 DLENTKRSQ---QAIKD---LNEFLKNNPNDQASKTLAQANKIOLHELSKVKHSIKEI 716  
 QY 634 --SSTLSQ---ITNAFDQNDGDKRGKCVVRTHNSHMLADERFIDPDKLSTIV 687  
 Db 717 DLENTKRSRQAKDNEFLKPNPNDQASKLQAAYENNGDLKVAENAEKIKLTNTQE 776  
 QY 688 D-----LEKVKITBGLDYSTIROTOTOPREMTTAGKAL---ISKSNUROGEKOK 735  
 Db 777 DHYKUGIRFKKVKYEHESFDP-OTIKLDPKHK-KALMKGIALMMLNKKALESFEX 834  
 QY 736 FLOBAQFFLGVRUDQDKAIAKAKALVTK-----ATKNGHILERSINKAV----- 780  
 Db 835 A1Q-----IDKNGTAYQKGABEKGMDQAFASPKNAVLNDQPKNPAVAKAGIVS 886  
 QY 781 -----LAVNNSAIKKANVRLKEBLDILTLIVE----- 808  
 Db 887 NNLGNFQKOSBEYLNFFNNAKAKPNEIATVNLISAFKPNKLESELETTINKAIDLNPEKER 946  
 QY 809 -----GKGPLAQATMVGVLKILPLPEYIYIINVYFDKGKULYALD-MSIT 857  
 Db 947 YLYKASINUKKVNQNALSYLSLVEKP-ENTSAVYIILAKAVEKGNSQALSTLECI 1005  
 QY 858 IGEGOKDAYEN-----PILNDEDENEYITLAVATLADYEGLYK 897  
 Db 1006 INKNNKALNLGILYKKEKNYQKAEIFFKAIN-SDEAKNLTATLIEIMDNTRAK 1063

QY 898 DLANSLDKIKATRQIPIAKYHRIGFOAIRMAAAEADRLLPKTPKGYLINEPNYRKQOM 957  
 Db 1064 DLAR-EYTKLKPMPNPEAL---HALGITEYENNNDQTUREL----IKKFPNPKNE- 1111  
 QY 958 EKULPKV 964  
 Db 1112 --NIKKI 1116

RESULT 15  
 T28677  
 C;Species: Plasmodium yoelii  
 C;Accession: T28677; C45521  
 R;Keen, J.; Sinha, K.; Holder, A.  
 Mol. Biochem. Parasitol. 65, 171-177, 1994  
 A;Title: A gene coding for a high molecular mass rhabdity protein of Plasmodium yoelii.  
 A;Cross-references: EMBL:27838; NID:9457145; PID:9457146; PIDN:AAA21304.1  
 R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
 Mol. Biochem. Parasitol. 42, 241-246, 1990  
 A;Title: Identification of the gene for a Plasmodium yoelii rhabdity protein. Multiple  
 A;Reference number: A45521; MUID:9101660; PMID:2270106  
 A;Accession: C45521  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 2131-2269 <KEE>  
 A;Cross-references: GB:M34283  
 Query Match 3.9%; Score 207; DB 2; Length 2269;  
 Best Local Similarity 18.8%; Prod. No. 0.16; 432; Indels 340; Gaps 58;  
 Matches 226; Conservative 206; Mismatches 98; Gaps 58;  
 QY 29 ERERSVKOBOTOSASEDDWFEEDNERKTKVSKENSTV-----DETVSDLFDGNSNNNSK 85  
 Db 989 DHDYNIKELKHKDKNSNGYKTEADQNKAIQKQKBLFEOKEEVILANKYKAVELKMKF 1048  
 QY 86 ESVSDPKQVK---AKPVEPOEASNSNDASKV-----EVPKDQDASKK----- 127  
 Db 1049 DKTKNDSKQKLTKEKTDARNYCTLESKGRKBBKMBKNEKHTEDEVANNDNSKNAITSTK 1108  
 QY 128 --ETLLETSTWEAKDVFVTGDTLVGFSKGINKLQTSHLVLPISHADQCTQLTQVASFAF 184  
 Db 1109 VSVEFPKUKLTKNEIRTKSD-----DCLEKETNDL-----RQISNL-- 1145  
 QY 185 TPDKKTAAEYTRSLGKPSRLDQKELIDBEGIINAYQJKTLPKLTGNGYKSGODAF 244  
 Db 1146 -----SITDQETKLTENG-----QLKTL-BELLSLKQKNTED----- 1180  
 QY 245 VDNKNTIAEVNPLSSELETSYDARHMSLQVKUPDNLYKIGELAFTPQDQIGKYLPRHL 304  
 Db 1181 -QKKEDEVN--SKK KNLENTRVNOHKKNVNEIGIVEKNEIATKIN--KNOIESTKLKPT 1235  
 QY 305 IKLARAFKSNRTOV---EFLG---SKLVIGEASFDQDNLRNVMULDGKLEKSEAF 358  
 Db 1236 IQHNTSENWANDLEGIDSDENLGKNTENGMNTYEEFIKSYNL---ITNLETVKESIT 1291  
 QY 359 GNPGDEHYNQVQVTRTQPHOLAMENTYVADPKSMLRATPDMDYKWLSDPTYQK 418  
 Db 1292 -----YNOI-----QN-KRDTQKLLKNNLENVNAKSYLDYK---EMF-- DR 1330  
 QY 419 SVTGSNKGQKVKRNKLTIPKONGI-----TTEBID-NAPRVDOSKTL--- 466  
 Db 1331 IVTHKKK-LATVNDNPKNEYSKNGFDNISNINTVGNSTBNSLNLINTQKEMTA 1389  
 QY 467 -----RKYDREI-----KLPSTIRKIGAFAPQSNLKSPEASEDLREIKEGAFMN 513  
 Db 1390 IVNNVYSYKAEABNFRNPQLANTLN-----KIKNSSGIDLSKJ-KIAISYLDs 1442

QY 514 RIGTDLKDOKL-----KIGDAAFHINHIIAYIWP--PSVOBGRSAFRONGALHMP- 564  
 Db 1443 K-----TDTLIPSPQKCTETYKISDSYSLILDKKSQELQK--KEQOTKLIPE 1494  
 QY 565 --IGNKVKIGMARFLNSKLESVNLSEOKLTLLEVQFSDNASEVULP-PMQTLIE 620  
 Db 1495 NRRLYKEKVQATNE--LRGTLSPDCKKKEKILSEVTLHHSNBLNKUSCNFQYDTILR 1551  
 QY 621 EA-----FCKNHLKEVKGGSTLSQTFNAFDONGDKRGK--KVVVRTHNSHMLADG 672  
 Db 1552 SSKYDQVKERSNNYKOEKGID----FNTMEEKENNDKVIELLENNDSEEN 1605  
 QY 712 FREMTAGKALLSKMRQGKQELQEAQFLG-----RVDLOKAIATA 755  
 QY 1673 ERFIDPDKUSSTM---VDEKV-LKTIIGID-----YSTLQTTQ 711  
 Db 1606 NNILOSKQKCLKELTNKFNAAEIKKDDKTIERKNNDLIDKLTETRKNOMLFHTTIAETLK- 1664  
 QY 757 EKALVTKKATKNGHLERS-----INKAVLAVNSAIIK 789  
 Db 1714 VTKSUDATNNNLIEKEFATOAKNLKFTIDSNTANALHNNKLOMWNSELH 1773  
 QY 790 KA--NYTRLEKEBLDILTDLVECKGPLAQATWQGVYL---LKTPLPLPBYIGNV-- 840  
 Db 1774 KSIESTKOLYKQHVKVQLNIGQ----INGKFDISKQFDNLQLQSELTANLND 1825  
 QY 841 -----YFDSKGKLIALMDSDTGRGQDQAYGNFTLNTYDDENEGHTIA-----V 885  
 Db 1826 LKEIGOKISDKKNEFLHNLNT-----PIPNTLKEIYHDTVKYKRDIEI 1872  
 QY 886 ATLADYEG---LYKDIINSLDKAI-----RQIPLAKHIGFOAIRMAAE 933  
 Db 1873 ENITSEBNENNTIYI-DTITKKEKTYOSILNFVTYENDSNTIKOH--TQDTNENDVSK 1928  
 QY 934 ADRLLPKPKGY---LNEVPNTRKKOMEKN-----LKFVDYKTPIF 971  
 Db 1929 IKESLKTIQSQBILANKINGIKAQFYNNNNNNNTISTISQVNDVKHISKDLTIE 1988  
 QY 972 NKALPNEKVDGRRAKHTINATNTNSWAVTPIRSEQOLHKSQDSVNLQPTSSKRNFIVE 1031  
 Db 1989 NELIEIQSLEDIKNTIEIRGANNNTNTNIVFQQNKIONNSNKEI--DDIQK 2045  
 QY 1032 ILGY 1035  
 Db 2046 TINY 2049

Search completed: April 27, 2004, 11:16:32  
 Job time : 34 secs



FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134, 001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064, 964  
; PRIOR FILING DATE: 1997-11-08  
; PRIORITY NUMBER: US 60/055, 779  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO: 3159  
; LENGTH: 10182  
; TYPE: PRT  
; ORGANISM: *Staphylococcus epidermidis*  
; US-09-134-001C-3159

Query Match 4.1%; Score 222; DB 4; Length 10182;  
Best Local Similarity 19.7%; Pred. No. 1.2e-05; Matches 236; Conservative 200; Mismatches 435; Indels 324; Gaps 58;  
Matches 236; Conservative 200; Mismatches 435; Indels 324; Gaps 58;

QY 331 GEASFODNNLRNVMGLPDGLEKIESEAFTGMPGDEHYNNQVULRTTRGQHQL----- 383  
Db 1584 -BAIOLANQKS-----NEARTALO-NERNSNNGQAQAKENGIRIBELWMPDAIK 1630  
QY 384 -----ATENTVNPDKSLRATP-----MDYTKMLEEDTYQKNSVTGFSNKL 428  
Db 1631 KSDAKOSIDNCKY-NEOSNTINTTPDADEEKOKALDKL-TAKDAGYNK 1678  
QY 429 OKVRVKNLEIPKQHNGITTEIGONAFR---NUFDQSKLRKDLEIK---LPST 479  
Db 1679 DQAOHQVSQVSDAKTEADITIONIQANVAKKPSARVELSK----FEDLKQINATPNA 1732  
QY 480 IRKGAFAFQSNNLKSFPEAS-----EDLEET-----KEGAFM-- 511  
Db 1733 TBEEKQDAIQRLNGKRDVEKVLINQDRRDNEVEQHKNIGHOLETHANPRTKSDALQEL 1792  
QY 512 -----NNRIGTDLKDK-----LIKIGDAAFH-----INHI 537  
Db 1793 QTKFISQTELINNKQATNEEKDEAKRMLTISKNTKTTINQATNQVNDNAKGMNEI 1852  
QY 538 YAIVLPSVQETIGRSAFRONGALHMP---GMVKVKTIGMA----- 576  
Db 1853 ATIPIAFTIKUDAKTAKKAEQVQTINGNNDATDEEKEARKLVEKAKLEAKSNITNS 1912  
QY 577 -----FLSNKLESV-NUSBOKQKJIEVQAFSDNALSFV---LPPNLQTREAF 623  
Db 1913 DTEREVNGAKINGLEKINNITOPSTQTKNAQEINOKAQBOLIQINTNTPTAEEQEA 1972  
QY 624 KR-----NLKXEVKGSSTLSQTFNAFDQNGDKRFGKVVVRTHNSHM 668  
Db 1973 NRVNAGLAQALONINNAHSTQEVNESTINSTATIKVQPNVTKPAAINSITQANQKT 2032  
QY 669 LADGERFTIDPDKLSSTMVDELEKVLKTELEGYSTLRTQTTQFREMTTACKALLSKNL 728  
Db 2033 LIGNDGNATDDEKEAKQKLVQTLNQBNQKIHSTQDNQDVNKAGAITATKLINANH 2092  
QY 729 ROGEKQKELQEAOPFLGRVDDKAIAKAKALVTKGATKNGHILERSINKAV----LA 782  
Db 2093 RQ-----DAINTLNLAEKKSDI---RANQDATTEEK---NTAQSIDDTLA 2134  
QY 783 YNNSAIKKANVERLEKELDLTLVGEKGPLAQATWQGVVILKTPPLPPEYYIGLNVYF 842  
Db 2135 QARNNINGANTVALVE-----NLEDQKQKQ----- 2165  
QY 843 DKSGKLTYALMDSTICRGQKQDAGNRPILNUDEDE-----GYHTLAVATLADYE 892  
Db 2166 STQTKQAKADIAQAGQQRSTIDQNQMATTEBKQEARLERINQETNGVNDRIOQALANQ 2225  
QY 893 GLYTIKDTINSSUDKIKARQIQLAKYHRLGTOAIRMABAADRDLRPLKTPK----- 943  
Db 2226 --VTDEBKNTLBTIRNVEPIVIVKPKANEI---IRKKAEOFTTLINQNDQATLEEKQIA 2279  
QY 944 -GYLNEVNPYRKOMERKLUKPVDPYKTPF-----KALPNEKUDGDKAAG 988  
Db 2280 LGKLEEVNEALMQVSQHNSNDVKIANNRATKISWHPETIIRKNAKOEIQAQSI 2339  
QY 989 HNINA---ETNN-----SVA-----VPIRSEQQLHKSQSDW-----IPQTS 1023  
Db 2340 DTINANNGSTNEEKSAAIDRVNVAKIDAINNNTNATTQLVNDAKNSGGTSISQPLSTA 2399  
QY 1024 SKNN 1027  
Db 2400 VKTN 2403  
  
RESULT 2  
US-09-134-001C-3159  
; Sequence 3159 Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Dorette-Stamm et al  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

QY 675 FID----PKLS-----TMIDLEKULK----ILRGLDYSRQPTQTO 711  
Db 9068 HTAENNTINHSDSISLANTLSKASLDNAMKOLRDTIESNSTSPVNSVNYINADKQIQT 9127  
QY 712 FREMTTAKGALLSKS-----NLRQEKQKPLQEAQPFJGRVLD 750  
Db 912 FDEALQASATSSKTSMPATIEEVIGLSQAYDFTKALNCE-QRIATEKSQDJKLTKGL 9186



APPLICANT: Harley, Calvin  
 APPLICANT: Andrews, William H.  
 TITLE OF INVENTION: No. 6093809el Telomerase  
 NUMBER OF SEQUENCES: 225  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 08/851,843A  
 FILING DATE: 06-MAY-1997  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 872 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein

US-08-851-843A-8

Query Match 3.7% Score 200; DB 3; Length 872;  
 Best Local Similarity 20.3%; Pct. No. 1.2e-05%; Matches 200; Conservative 143; Mismatches 350; Indels 294; Gaps 48;

QY 14 TTWSVVTISQEVGLE-----REESVK-----QEQTQSASEDDFEEDNERK- 55  
 DB 17 TNIDFVQNLLEVYKSQLTHYKUQQQTHGEDLKILKEPNQDGNSGNDDEENNSNQ 76  
 QY 56 -----TNUSKENSTVBTVSULPSDONSNNSSKTESVSDPKQVKAKPEVTOASN 108  
 DB 77 QELLRRVVAQIKQQVOLIKKVSKVKEKDNLN-----EDEN 111

QY 109 SSNDASKKEVPKDGTASKETLTSTWRA--KDFVTRGDTLUVFSKSG----- 154  
 DB 112 KNGLSEOCV-----KEQLRTRTEBVQKVQMLVENDYQOLDLNESSGGHRHRRTDY 164

QY 155 -INKLSQTSHLVLPNSHADGQTQVASAFAFDKKAIAEVTSRIGENGKSRDIDQK 213

DB 155 DTKEKWEFISH-----DOKNTVSIYA--NOKTSTCWMILKDYFENKNYDHLVNSIN 211

QY 214 EIDEGEF--NAYQLTKLITPGYKSTGQDAFDVNK-----NAEV- 253  
 DB 212 RLETEAEEVAFDPSQTSKRLT-NNSYQTVNIDYFNEDNNCLLALLRFLPLSERFNTLN 270

QY 254 -----NLPESETTSIDSYAFAHMSLKVQLPDLNKLVIGSLAFFDNQGK----- 297

RESULT 5  
 US-08-851-843A-54  
 ; Sequence 54, Application US/08851843A  
 ; Patent No. 6093809  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Czech, Thomas R.  
 ; APPLICANT: Lingner, Joachim  
 ; APPLICANT: Nakamura, Toru  
 ; APPLICANT: Chapman, Karen B.  
 ; APPLICANT: Morin, Greg B.  
 ; APPLICANT: Harley, Calvin  
 ; APPLICANT: Andrews, William H.  
 ; TITLE OF INVENTION: No. 6093809el Telomerase  
 ; NUMBER OF SEQUENCES: 225  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States of America  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

Db 271 SSYTRNQNFKEKELLETFATVFSRSHLOGIHLQVPCAEQYLVNSSSQISVKSQDSQ 330  
 QY 298 LY-----LPHLILKABRAKSNRQTVERIGSKLXVIGAS 334  
 Db 331 VYSFSTDQKUDTMKQDQYFKEFPR-LTHVSQALPVSATNAVENLVLKKVKA- 388  
 QY 335 FQDNNILRNMLPDSLEKIESEAFTGNGPGEHVNQVVRTRTSQNPOLATENTVYVPDF 394  
 Db 389 ---NINLYSIP-----TQFNPDPYFVNLQHICLEGFLEPNLTQKL---E 428  
 QY 395 SLMATPDMQYTKMLEDFTY-CQNSYTGFSNG--LOKVRNKNLIPKOHNGITIT 449  
 Db 429 NLLLSIKQSKNKLKEPLRNINFTYVAQETSRQKQATTIKLNKNNKNOETPEKDEPS 488  
 QY 450 EIGDNAPRNVDQSKTIRKYLEDKLPSI--IRKIGAFQSNLNSPEAS-EDLIE 504  
 Db 489 E-SRGMKRFDHLSBLTEBDF-SVNLQTOETYDSIHLKLRLTNKPLKSYTMEK 546  
 QY 505 IKGAF-MNMRIGTL-DJKDQLIKGDAFHNNIYAVLPESY-----QEI-G 550  
 Db 547 SKMDTFFIDIANIYETLANNLKCSTYNSP--HGNISYELNKDSTFYKPLTQNLQELQHA 604  
 QY 551 RSAFFQNGALHLMFEGNKVKTIGENAFSLNKLESY-----NISSEQQKLTIFWQAFS 602  
 Db 555 KTFKQN----EFQFNNVKS--AKIESLIESJEDIDSICKSIAASCQNLQV----- 650  
 QY 603 DNALEBVVLPPNLOQTREAFAKRNLI--KEVKSSTLQQTEN-FDONGDKRPGKKV 658  
 Db 651 -NITISLIVPNIQ--KNFNPKNPILFFQPEQTKRNLENVINCILDO----- 695  
 QY 659 VVRTHNMSMLADGSRFIPDPKLSTMDVLEKUKLICEDYDSTRQTRQ----- 709  
 Db 696 -----HINNSISBPFLEKNNKTKAFLKRYVLLQY--LDTYTKLFTLQLQPLBNQV 745  
 QY 710 --TQFEMTITAGKALIJKSNSRQGKQK-----FLQAQFEGRVIDLKAIA--KA 756  
 Db 746 INQQLBELTWSH--VHKQWENKMQKAFYEPICLFFKESQTLQIDPNTQWSDSI 801  
 QY 757 EKALVYKKAATKNGHLERSTINKAVLAVNSAIKKANVKRLEKEIDLTDLVEGKQPLAQ 816  
 Db 802 KKILSEISSEKHHVHLRNLNSQ---SSSIKSEN---EELQBLIKACDEKSYLVKA 851  
 QY 817 TMVQGVYLLKPLPPEYVIGLNVYFD 843  
 Db 852 -----YYKFLCLP----TGTYYD 866

SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/851,843A  
 FILING DATE: 06-MAY-1997  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002930US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 54:  
 SEQENCE CHARACTERISTICS:  
 LENGTH: 872 amino acids  
 TYPE: amino acid  
 STRANEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 US-08-851-843A-54

Query Match 3.7%; Score 200; DB 3; Length 872;  
 Best Local Similarity 20.3%; Pred. No. 1.2e-05; Gaps 48;  
 Matches 200; Conservative 143; Mismatches 350; Indels 294; Gaps 48;  
 QY 14 TTVSVVTVSQEVYGLE-----REPSVK-----QEQTQASASEDWFEDNEERK- 55  
 QY 17 TNLDLFLVLOMLEVYKQSQIHEVKHTQQQKIREDKDILKLFKNQDQDONGSDDEENNSKQ 76  
 QY 56 -----TIVSKERNTVDTEDTIVSDFSDGNSNNSSKTSVSVSPKQVKRPEVTOASN 108  
 QY 77 QELLRRVQVQKQOQVOLIKVKGSKVEKDLNL-----EDEN 111  
 Db 109 SSNDASKVYVVKPQQTASKETLESTWAA--KDFVTRGDTVGEFSKG----- 154  
 Db 112 KKGGLSEQVW-----KBRQRLTTEVKYQMLVENDYQDLDNESGHRHRRRTY 164  
 QY 155 -INKLSQLTSVHLVLP SHAADGTQVASHAFTPDKTTAIAEYTSRIGENGPSPRDIQK 213  
 Db 165 DTEKWFELIH-----DKQTVYVIA--NQKTSQCMWLKDYNENNNYHHLNUSN 211  
 QY 214 EITDEGEGTF--NAYVQLTUUTIPNGYKSGQDADYVDNK-----NIAEV- 253  
 QY 212 RLETEAEEFVAFDDFSQTIKLT-NNSYQTNDVNFNNICILALLRFLISLERNLINR 270  
 Db 254 -----NLBESLETIUSDYVAFHMSLQKQVLPDNLKRVIGLAEFFNQIGCK----- 297  
 Db 271 SSYTRNQYINFEKIGELLETFIAVVFQSHRLQGHLQVPEAFQAFQVLYNNSQSISVRSQDQ 330  
 QY 298 LY-----LPRHLIKLKAERAFKSNRQIQTVERFLGSKLKVIGEAS 334  
 Db 331 VYSFSTDLKLVDTNKVQDYFKFQFPR-LTHVSOQATPVSAVENNLVLLCKVKA- 388  
 QY 335 FODNNLRNWLDPGKEKIESEAFTIGNPQGDHYNNOVLRTRTGQPHQATENTYVPDFK 394  
 QY 389 -----NLNLVSP-----TOPNFDPYFVNLQHILKFGLEPNLTQKLU-----E 428  
 QY 395 SLWRAATPDMYDITKMLEDF-TY--QKNSVIGPSNKG--IQRVURRNKLSPKQHNGITIT 449  
 Db 429 NLLISIKQSKNLKFLRNLFTIVAOETRSRQLKQATTIKLKLKQKQSETPKDETS 488

RESULT 6  
 US-08-974-549A-221  
 Sequence: US/08974549A  
 ; Sequence 221, Application US/08974549A  
 ; Patent NO. 6165178  
 GENERAL INFORMATION:  
 APPLICANT: Cech, Thomas R.  
 APPLICANT: Lingner, Joachim  
 APPLICANT: Nakamura, Toru  
 APPLICANT: Chapman, Karen B.  
 APPLICANT: Morin, Gregg B.  
 APPLICANT: Harley, Calvin B.  
 APPLICANT: Andrews, William H.  
 TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
 NUMBER OF SEQUENCES: 727  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 08/974,549A  
 FILING DATE: 19-NOV-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/911,312  
 FILING DATE: 14-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,951  
 FILING DATE: 14-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US97/17885  
 FILING DATE: 01-OCT-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph Ted  
 REGISTRATION NUMBER: 36'429  
 REFERENCE/DOCKET NUMBER: 015389-002610US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 221:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 872 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-974-549A-221

Query Match 3-7%; Score 200; DB 3; length 872;  
 Best Local Similarity 20.3%; Pred. No. 1.2e-05; Indels 294; Gaps 48;  
 Matches 200; Conservative 143; Mismatches 350;

QY 14 ITIVSVVIVYQSYVIGLE-----REESVK-----QEQTQSASEDDWFEEDNERK- 55  
 Db 17 TNIIDFVULQVLETYKSCIEHYKQHQDQDQKFLKQKPPDKQDQGHSNDDDEENNSKQ 76  
 QY 56 -----TIVSKENSTVDETDVSLFLSDGGNSNSSSKTESVVSDDPKVQVAKPEVTOBASN 108  
 Db 77 QELLRRNQIKQVQVQLIKVSKVKEADLN-----EDEN 111  
 QY 109 SSNDASKVKEVVKPQDTASKETLETSTWEA---KDFVTRGDTINGFSKG----- 154  
 Db 112 KNGLSLBQQV-----KEEQRTTIEBQVKIONLYVNMDIQLDNEGGRRHRRETDY 164  
 QY 155 -INKLSTTSHLVLUPSHADGQIQTQVQASFAFPDKTIAFVTSRSLRGENGKPSRLDQK 213  
 Db 165 DTBKWEFISH-----DOKNYVSYA---NQKTYCWM----- 211  
 QY 214 EIIDEGIF---NAYQITKLITPQNGIKSIGQDAFDVNDK-----NIAEV- 253  
 Db 212 RLETEAIFIYAFEDFESQQTKL-NNSYQVNTLNFNDNLCTIALRPLSLERFNINR 270  
 QY 254 -----NLPESLTISDYAFAMNSLKVQVQLPDKLKVIGELAFFFQDQIGK----- 297  
 Db 271 SSYTRNQYNFKEKIGELLETIFAVFVPSHRLQGHLQVFCEARQYLNSSQLSVKSQSQLQ 330  
 QY 298 LY-----LPRFLKLAERAKFSNITQTFELFGSKLKVIGEAS 334  
 Db 331 VASFSTDIKLVDTNKVQDIFYKFLQEFER-LTHVQSQQIPVPSATNAVINLNVLKKYKA- 388  
 QY 335 FQDNTRNQVMLPDGLEKIESEAFTGNCPCDEHYNQVWTRGTSQNPOLATENTVNDK 394  
 Db 389 ---NLNLVSI-----TQNFNDFYFVNLOHLKLFQLEPNLITKQK-----E 428  
 QY 395 SLMRATPDMYTKLLEDFP-TY-OKNSVIGFSNKG- LQKPRRKKRKLIPKOHNGIT 449

RESULT 7  
 US-08-854-050-8  
 ; Sequence 8, Application US/08854050  
 ; Patent No. 6261836  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: Czech, Thomas R.  
 ; APPLICANT: Lingner, Joachim  
 ; APPLICANT: Nakamura, Toru  
 ; APPLICANT: Chapman, Karen B.  
 ; APPLICANT: Morin, Gregg B.  
 ; APPLICANT: Harley, Calvin  
 ; APPLICANT: Andrews, William H.  
 ; TITLE OF INVENTION: No. 6261836el Telomerase  
 ; NUMBER OF SEQUENCES: 225  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States of America  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ;  
 ; COMPUTER TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/854,050  
 ; FILING DATE: 09-MAY-1997  
 ; CLASSIFICATION: 536  
 ;  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/851,843  
 ; FILING DATE: 06-MAY-1997  
 ; CLASSIFICATION: 536  
 ;  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/846,017  
 ; FILING DATE: 25-APR-1997  
 ; CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724, 419  
 FILING DATE: 18-APR-1997  
 CLASSIFICATION: 536  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724, 643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36, 429  
 REFERENCE/DOCKET NUMBER: 015389-002930US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 872 amino acids  
 TYPE: amino acid  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 US-08-854-050-8

Query Match 3.7%; Score 200; DB 3; Length 872;  
 Best local Similarity 20.3%; Pred. No. 1.2e-05; Mismatches 200; Conservative 143; MisMatches 350; Indels 294; Gaps 48; QY

QY 14 |TVSVVTVYSPBVYGLR-----REBSVVK-----QEQOTOSASEDDWFEEDNERK- 55  
 Db 17 TNLDPLVQNLLEVYKSIQHETKQQQIKEDDILKIKFQNDQDGNSGNDDDEENNSNQ 76  
 Qy 56 -----TIVSKENSTVDEWVSDLIFSPDGNSNNSSSKTESTVSDPKQVPRAKPEVQEAQN 108  
 Db 77 QELLRRVNVQKQVQVQKVKVGSVKERDNLN-----EDIN 111  
 Qy 109 SSNDASKVEVPKQD TASKESTLSTWEA---KDVVTGDTLVGRGSKG----- 154  
 Db 112 KKGKGSEQV-----KQSLRITTEQYKQNLVENDYQDLDNESGHRHRETY 164  
 Qy 155 -INKLSQTSVHLVPLSHAADGTOVASYFAPTPDKCTTAAEYTSRSLIGENKPSRUDICK 213  
 Db 165 DTEKPEFISH-----DOKNVSIA---NOKTSCWMLKDYFNKNYDHANSVN 211  
 Qy 214 EIIDGEIFF-----NAVOLTKLITPQYKSTQGDAFYDNK----- 253  
 Db 212 RLETERAEFYAPDDESQTIKLT-NNSYQTVNTVNDVFDNNCTTALLRFLSLERFNLNIR 270  
 Qy 254 -----NLPSIETISDYAFAHMSIKQVKUPDNLKVIGELAFFNOIGSK----- 297  
 Db 271 SSYTRNQYNEFKIGISLLETIFAWFVSHRQHIGHQVPRAFQVILVNSQSISVQDSQ 330  
 Qy 298 LY-----LPRHILKLAERAFKSNRIOVVEFGSKLVIGEAS 334  
 Db 331 VYSFSTDLKVLNDTNUQDYKFKLQEFPR LTHVSOQAIPIVSATAVENAVNLKVKHKA- 388  
 Qy 335 FODNNLRNMLPDPGLKIESEAFTONPGDHEHNNQVUVRTGCPHOMHQATENTYVNPK 394  
 Db 389 -----NLNLVSI-----TOFNFDYFVNQLOHLKLFGLEPNLITQKL-----E 428  
 Qy 395 SLWRATPDMYTKLLEDF-TY-OKNSVYGSNG--LOKVRANKNLIPKQNGIT 449  
 Db 429 NLLSITKQSKNKLFLRNFNTYVQAOTSRSKQIKAATTKNNKNNQSRPTEDETPS 488  
 Qy 450 BIGDNAPRNVDQFOSKLRLKDLBEEKLPLST--IRKIGAFAFQSNLJKFBEAS-EDLSE 504  
 Db 489 E-STSGMKFFDHLSELTELEDF-SYNLQAOETVSLHKLIRSLNKFPLSKYEMK 546  
 Qy 505 IKGAF--MNRIGTL-DLKOKLIKIGDAAFHINHIVATVLPESV-----QEI-G 550  
 Db 547 SKMDTIPIDLNKVIETTNNLKRCSVNTSPL-HGNISYELINKDSTPYKFLTINQBLQHA 604  
 Qy 551 RSAFRONGALHLMFTGNVKTIGENAFLSNKEV-----NLSEQQKLTIEVQAFS 602

RESULT 8  
 US-08-854-050-54  
 Sequence 54, Application US/08854050  
 Patent No. 6261836  
 GENERAL INFORMATION:  
 APPLICANT: Cech, Thomas R.  
 APPLICANT: Lingner, Joachim  
 APPLICANT: Nakamura, Toru  
 APPLICANT: Chapman, Karen B.  
 APPLICANT: Morin, Gregg B.  
 APPLICANT: Harley, Calvin  
 APPLICANT: Andrews, William H.  
 TIME OF INVENTION: No. 6261836el Telomerase  
 NUMBER OF SEQUENCES: 225  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 ZIP: 94111  
 COUNTRY: United States of America  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 COMPUTER: IBM PC compatible  
 SOFTWARE: Patient in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/854, 050  
 FILING DATE: 09-MAY-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/851, 843  
 FILING DATE: 05-MAY-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846, 017  
 FILING DATE: 25-APR-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724, 419  
 FILING DATE: 18-APR-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724, 643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36, 429

REFERENCE/DOCKET NUMBER: 015389-002930US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 54:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 872 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 US-08-854-050-54

Query Match Similarity 3.7%; Score 200; DB 3; length 872;

Best Local Similarity 20.3%; Pred. No. 1.2e-05; Matches 200; Conservative 143; Mismatches 350; Indels 294; Gaps 48;

Qy

14

TRVSIVVTVSQRYGLE

-----REESUK







Db 331 VYSFSTDLKLUVDTNKVQDYPKFLQBPFR-LTHVSQATIVSATNAYENVNVLKVKH- 388  
 Qy 335 FODNNLRNMLPDGLEKISEAFTGPNPGIHEHNNNOVVLRTGPNPHOLATENTVYNDK 394  
 Db 389 ----NLLNJSIP----- 428  
 Qy 395 SLWRATPDMDYTKLEEDF-TY--OKNSTGFSNKG--LQKVRNKNLEIPKGHNGITIT 449  
 Db 429 NLLISIKQSKNLKEPLNFTYVAQETSKQIQLQATTKNLKNNKGQEPETPEKDEPS 488  
 Qy 450 EIGONAFRNVDFOSETKLRYDLERKPLT--IRKIGAFRQNNJXPEAS--EDIE 504  
 Db 489 E-STSGMKFFDHSBLTEBDF-SVNLOTOQETYDSLHKLINSTNLKFKLSSYKEMR 546  
 Qy 505 IKEGAP--MNARIGIL-DIKOKLTKIGDAFHINHIVATVLPESV-----QEI--G 550  
 Db 547 SKMOTFIDKVNLYETLNKCSNLSN--HEMISYEN-TNKSTPXYKEKLTNOELQH 604  
 Qy 551 RSAFRONGALHMTGKNTGKTTGEMAFNSNLKSV-----NLSEBGOKLKTIEVGAFS 602  
 Db 605 KYTFKQN---EFQFNNTKS--AKIESSLESLEDIDSCLCKSIAASKNLQW----- 650  
 Qy 603 DNALSERVVPNPNLTIREAFKRNHL--KEVGSSTS-SQITENA-FQONDGDKRFGKX 658  
 Db 651 -NLLASLILPFNNIQ---KNPKNPKNLFFRQFQEOKLNENVSINCILQD----- 695  
 Qy 659 VWRTHNNSHMLADGEREIDPDKUSSTMDLEKULKLTGFLDSTLRTQ----- 709  
 Db 695 -----HITNSTSEBLRKKKKAFTKRYLYQY--LDTKLKFLQDOLPENOVY 745  
 Qy 710 --TOPREMTTAGKALLSKNSLROGEKOK-----FQOBAQFFGRVLDKIA--KA 756  
 Db 746 INQOQELLELTVSE---VHKQWENHKQKAFYFCEFIMESQSTQIQLIDFDQNTVSDDI 801  
 Qy 757 EKALVTKKMTKNGHLLERSINKAVTAYNNSAIKANVKELEKFLDUDLVLGKGPAAQQA 816  
 Db 802 KKULESISSESKYHYLRLNPSQ---SSSLIKSEN---ERIQELIKADEBKGVLVKA 851  
 Qy 817 TMVQGVYLLKTPPLPEVYIGLNYTFD 843  
 Db 852 -----YXKFPICL-----TGTYYD 866

RESULT 13

US-09-595-684B-31  
 ; Sequence 31: Application US/09595684B  
 ; Patient No. 654766  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Ohashi, Cara  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Vaisberg, Eugeni  
 ; APPLICANT: Wood, Kenneth  
 ; APPLICANT: Yu, Ming  
 ; TITLE OF INVENTION: Human kinesins and methods of producing  
 ; TITLE OF INVENTION: and purifying human kinesins  
 ; FILE REFERENCE: Cycop036  
 ; CURRENT APPLICATION NUMBER: US/09/595,684B  
 ; CURRENT FILING DATE: 2002-06-24  
 ; PRIORITY APPLICATION NUMBER: 09/295,612  
 ; PRIOR FILING DATE: 2000-04-20  
 ; NUMBER OF SEQ ID NOS: 105  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 31  
 ; LENGTH: 2662  
 ; TYPE: PRT  
 ; ORGANISM: Human

US-09-595-684B-31

Query Match 3.6%; Score 192.5; DB 4; Length 2662;  
 Best Local Similarity 20.0%; Pred. No. 0.0002; Matches 433; Indels 365; Gaps 59;

Qy 21 YSOEVYGLEREEVKQEBOTQASEDDWEEDN--ERKINVSKENSTVDETVS----- 70  
 Db 507 YDNLVLVDYEQLRERKEKEMELKEKEKNDDEFALERKTKKQDEMQLIHEISNKLVYHR 566  
 Qy 71 DLSGDGNNSSSKTTESVSDPKQVPRKPEV-TQASNSNDAS---KVEVKPQDTAS 125  
 Db 567 EVINQDLENLSSKVELRKEQDQKIKLQEYIDSQKLENKNDLISYLESIDPKQ--- 622  
 Qy 126 KRETL--ETSTWEAK--DFVTRGDTIVGFSKSGINLQSQISHLVLSHAAGTQLOTOVA 180  
 Db 623 MKQHFLDABTVALDKRBSAFLRSENLE--LKEKMKELATTY---KOMENDIQLOQ-- 673  
 Db 181 SPAFTPDKKTALAEYTSRULGENGKPSRUDQKETID-EGEFNAY-QLTKLMTIPNGYS 238  
 Db 674 ----- 701  
 Qy 239 IGGDAFVNKNIAEVNIPESLETTSYAFAHMSIKQVLPDNKURVGBLAFFNQGKL 298  
 Db 702 ---SLIDGK---VPKDLG--- 733  
 Qy 239 -----YLPRLHTKLAERAFPKSNRIOVE----- 392  
 Db 794 VMLPDGLKIEBAAFTG-----PEDEHVNQVWLRTRCPHOLATENTVYNP----- 392  
 Qy 343 VMLPDGLKIEBAAFTG-----PEDEHVNQVWLRTRCPHOLATENTVYNP----- 392  
 Db 794 LLEBKGTRKDDIATTSQSYKSTDFOEKFKTLMDFBQKVKYVLENERMNBQIVNLK 853  
 Qy 393 --DKSMURATDMDY-----TRKLEPFTYQSVTGSNKG--LQKRNANLKE 438  
 Db 854 AQKEDSSLGKELSYKTIQELQEKTRVQESINERMPQKQLENRISPLQTEREKTU 913  
 Qy 439 IPKQHNGI---TIT----- 472  
 Db 914 TEKLQQTLEEVKLTQDQKDLQKQESQIERDQKSDIHDTWMMNIDQQL--RMALE 971  
 Db 473 EIKLUPSTRKIGAFAGSNNLKSPEASDDELEKKEGFMNWRIGTL--DLKOKLTKITD- 529  
 Db 972 SLKHQETI-----NTLK---SKISEEVSRNHLMEENTGETKDEFOQKMGVSDKK 1017  
 Qy 530 --AAFHINHNTAIVLIPESVQIGRASFRONGALHMEIGNKXTKGEMAFNSKJESVN 586  
 Db 1018 QDLEAKNQTPLTADVKONEITBQORKIT----- 1061  
 Qy 587 LSBQKQKLT-----IEVQAFSDNALSPEVVLPPNLQTRE-BAFKRNHL--KEVKG-ASTLS 638  
 Db 1062 TAEKEQKTDLKENIEMTENIENEEZER--LIGEELKQKQELVQEQKNAIKKEGELRPTCD 1119  
 Qy 639 QITFNAFDQNDGDKRFGKVVTRTHNNSHMLADGEREIDPDKLSTMVDELKVKLTG 698  
 Db 1120 RLA-----EVEEKLKEVSQQLQEQKQQLIN--VQEMSEMOKKINIEEN 1161  
 Qy 699 LDYSLTRTOTTOPREMTTAGKALLSKNSLQGE-----KOKFLRQA--FELGVLD 750  
 Db 1162 LKNEL--KNNKEITLHEMETERLELACKLNENYEVVISITKERVKVLQKSPETERDHLR 1219  
 Qy 751 KATAKA-KATVTKKATKNGH-----LESN-KAVLANNSAKKA----- 791  
 Db 1220 GYTRETATGLQKEKLT----- 1279  
 Qy 792 -----NVKRLK-----BLDLDLVLGKGPLAQATWQGVYLLKTPPLPLP 832  
 Db 1280 EIPVLRHEQELIPPNVKVSETOETMNELELT-----QSTTKDSTLARIME- 1328  
 Qy 833 EYVIGLAVYFDKSKGKLYA-----DMSDTGEG----- 868  
 Db 1329 --RIRLNKEFQESQETKSLIKERDNLKTIKEALEVKHDQKHEHBTLAKIQESQKQ 1386  
 Qy 869 P1LNVE-DNSGQHTLAVATADYECILYKQILNSLUDKTAIRQPLAKHRL-GIFQA 926  
 Db 1387 OSLNMMKEDNE---TTKUVESEMEEPKPO---SALRUE-TEMGILSK--RLOSHDE 1435

RESULT 14

US-09-107-532A-3959

; Sequence 3959, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

APPLICANT: Lynn A. Doucette-Stamm and David Bush

NUMBER OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/0551571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)823-5007

TELEFAX: (781)821-8277

INFORMATION FOR SEQ ID NO: 3959:

SEQUENCE CHARACTERISTICS:

LENGTH: 1221 amino acids

TYPE: amino acid

TOPLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc-feature

LOCATION: (B) LOCATION 1..1221

SEQUENCE DESCRIPTION: SEQ ID NO: 3959:

US-09-107-532A-3959

Query Match 3.5%; Score 189; DB 4; Length 1221;

Best Local Similarity 19.0%; Pred. No. 0.00012; Gaps 61;

Matches 228; Conservative 174; Mismatches 403; Indels 394;

QY 4 KHKTIAAL-TIVSVVTTQSYEV--GLEREESVKQEQTQSASDDWPEEDENERKTVS 59

QY 14 KKUKLISIALFSLTMSTARQSVIATVSGTSTKAWVEESGDKVAIDDKVKNRERAKTIVQ 73

QY 60 KENS---TVDTIVSDFSDGNS--NNSSSKTESV--VSDPKQ--VPAKCFVTDORASN 108

Db 1436 MKSVAKEKDDLE----ORLOEVLOSESDSLKENIK----EVAKHLTE---EELK 1479

Db 987 KGSH--NINATNNNSVATPIRSE-----QOLHKSQSDNLPQT 1022

Db 1480 VAHCCLKEQESTINELRVLNLSKEKETRISTICQKLAINDKLUQNQK1QEIYEKEQNLQK1 1539

Db 1023 SSKNNPFIYI 1032

Db 1540 SEVQDNVNL 1549

Db 74 GADGLMTWVVDENDPEPKLEAESPPLKQHQENGAVEVGVDERGQEKADKANEKVKNFDDI 133

Db 109 SSENDASKEV--PK----ODTASKEETLETSTWAKDFV--TRDTLWFRSKSGNK 157

Db 134 DNDISLQLSALNDPKPFPSSEDELNLKQTKNTESEALMKTR---GFSORIE 188

Db 158 USQTSHLV---LPSHAADTQLTQWASPARTDPDKTALARYTSRIGENG---KPS 206

Db 189 USNNLNUVPKRAKVSRHALVNSRASOSETTRNPPNGPYPWVYLIR---SNGSYNTENAI 246

QY 207 RIDOKEIL-----DEGEIKNAYQKLTUINGYKIG----- 240

Db 247 ARTIDGKEARFLQAGVQATNTTGTGSTGGLNSQRENNISLASYKELGWANNDACKQ 306

QY 241 ---QDAFDVNKNIAEVNLPLPSELETSVD-----AFARMSLKQVKLPDNKVI 284

Db 307 WIATOMLWENSN---EKPPSITVPRWKNGATKMDLYARANAIYRUVKNDISKIV 362

QY 285 --GELAFF-----DNOI-----GGKLUPLRHEKLKLAERAKFSNR1QTYEFL--- 323

Db 363 DGDGKATPKVNTYGDHNLHVKVSVGGGKATISNGITVDTGATGK1-TWDIMKSHP 421

QY 324 -----GSKLKVIGRASFQDNLR-NVMLPGLC-KIEST-----AP 357

Db 422 SDAHPTTVWSNGSYOKLUTSFYIYSTVWVTKRKGTVVVKVUDSTKKPISGATERFSY 481

QY 358 TGNPGDEHYNQVULTRTGONPHOLATE---NTVNPDPDSLWRATPDMDYTKWLEDF 413

Db 482 SGKTKD-----VTTGNGKAKLTERLKGATVVK-----VKEIKAPN 517

QY 414 TYQK1---SVTGFNSKGLOKVRNKNLEFRKHNGITITBGD---NAFRNDF---- 461

Db 518 TYQDLSSEFSITIVKQNQVNTTTRNK---KSTGSEVIEKIGDGLGLENVFTIINS 572

Db 462 QSKTR-----KDYLIHKLPS--TIRKIGAFAFONNLK-SFEASDEELEKEGA 509

Db 573 DNKVVKDNLKDTSNGKIKUDLOFGSKYAYEVQGTVGDFGKYNFELSQDTPFTK-- 629

QY 510 FMNRIGTFLDIDKLUKIGDAFHNFHAYAVLPLSYQERGSAFRQNGALHLMFTGKV 569

Db 630 -----PAKVKYNTIIEVSISKEVSDMNEQSVTKNN----- 659

QY 570 KTIGEMAFLSNKLESVNL--SEQKQKTEIWQAFSDNALSEVVLPPNLOQTIREEAFKRNH 627

Db 660 WSADPNQDLYQDITGTLKQKSGERNVOSFTIEGFVN-----KVNQVSKSEV--- 707

QY 628 LKEVKGSSITLSQITFNAPDONGDQDKRFGKVVVRTHNSHMLANDGERFTIDPDKLUSTMV 687

Db 708 ---VGSTNVT---NSFDINNDT-----NNGKVIAKRTSVLNNDFNPKSY 747

QY 688 DLEKVKLKTIGGLDYSTLROQTQFREMTAGKALLSKNLRDSEBQKQLEQAFGLGRV 747

Db 748 NLRTMKI-----KGTIENTEKQNL----- 769

QY 748 DLDKIAKAKLAKVYTTKATKNGHLLERSINAKVLAYNNSAIKANVRL-EKEUDLTL 806

Db 770 ---TIVKNSKVTVGSKLRSKA---SSNDVETVLYS---RKVTINHIDEKDQHLLQD 817

QY 807 VEKG--GPAQATMVQGVYLLKIPILPPIYYIGLAVTFKSGKLIYALDMSTTGBGQKD 864

Db 818 IDYKQDGETYE-----YKPRTDI-----FDKDGNNY---KSDVHKGKD 854

QY 865 AYGNPILNVDNEGYHTLAVATLADYEGLYKIDNLNSLDDKUAIROFLPLAKYHRLGIF 924

Db 855 --GKDIVL---NTPYH1-----PULSVNDR1---QDTR----- 882

QY 925 QALRNNAAEADRLPKTPKGY-NEVNPYTKOMEKLNKLPDVYKPIFENKALPKNEVKDGD 983

Db 883 -ATVN-----GILPKIPELSKKEEKELEKVKVWIKTID--IDNNKVVYD 925

QY 984 RAAGKHNNAETNSVAVTPIREBQOLHKSOSDYNLPPTSSKONFYE---ILGVS 1037

Db 926 ENFKYKERNHMSLDLPIEYLTNNKKINNS-VDILLVNDPKNKFETETKDLRTIGYS 983

RESULT 15  
US-08-592-126-148  
; Sequence 148, Application US/08592126  
; Patent No. 5821091

GENERAL INFORMATION:  
APPLICANT: Gregory Dolganov  
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory  
TITLE OF INVENTION: Polypeptides  
NUMBER OF SEQUNCESES: 151

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dellinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,126  
FILING DATE: 2003-03-11  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles R.  
REGISTRATION NUMBER: 38,615  
REFERENCE DOCKET NUMBER: 4600-0111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1312 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54  
US-08-592-126-148

Query Match 3.5%; score 188; DB 2; Length 1312;  
Best Local Similarity 19.2%; Pred. No. 0.0015; Mismatches 220; Conservative 185; Indels 296; Gaps 50; Matches 220;

Qy 18 VVTISQVYGL-REESVKQFQQTOSASEDDWFEDNEERKTKVNSKENNSTVETVSDLF-S 74  
Db 241 VKSYENELDPLKNRKEHNLKSLTKMLNKLDSRKKQMKRMEKVFGT 300

Qy 75 DGNNSNSSSKTESVVDPKQPKPFRVTOAASNNSNDASKVKEVPKDQKTTST 134  
Db 301 DQLQNLKHNFIORTV--REKERKLVCHREBLEKLNEKSLNQEKSELLVEQGRLOQA 357

Qy 135 WEAKDFVTRGDTLVGFSKGNGINKLSCOTSHLVLPSHAADGTLTQVASFATPDKTAIE 194  
Db 358 DRHQEHIRARDSLI-----QSLATOL-ELDGFPERGPFSERQIN 395

Qy 195 VTSRIGRNGKISRLDIDQKEI-DEGRIFINAVQLTQKLTIPNQYKSLQDADVNKAEN 254  
Db 396 FHKLVRBR-----OEGHAKTANOL-----MNDFAEKETIKOKQ 428

Qy 255 LPE-----SLETISYKAFAMSLKQVPLDKVIGLAFPDNQIGKYLPRULIKIA 308  
Db 429 IDEIRDKKTFGIGRILKSKKQNL-KNQVY-ELQLEGSDRRIEDQSLIKE 485

Qy 309 BRAFKENRITQTEFLGSKLVY--GEASFOODNNLRNMLPDGLERKIESRAFTGNPGDEH 365  
Db 486 RELSKERKNSWETL-KMEVISLQNEKADRTL-----KLDQEM-----EQ 527

Qy 366 YNNQVULTRTGQNPOLATNTYNDPKSLW-RATPDMDY-TKMLEDPFTYQKSVT 421  
Db 528 LNHHTTRTQEMLTKDQKADKEQKTKRSRHSDELTSLGYPFNKQLDWLSHKSKEI 587

Db 422 GFS-----NGKQKRNK--NEIPIKNGHTITEQDN-----RNDFOQTLRK 468  
Db 588 NOTRDRHLAKLNLKELASSEQNHNINNELKREEQLOSSYE-DKLFUVCGSDPFS 640

Qy 469 YDL---EEIKLUPSTIRKI--GAFAQSNNL-----KSFBASEDIEIKE 507  
Db 641 -DIDRKEIEKSSKQDAMLAGATAVSOTIOLTDENQSCCPVCQRFQTEALQEY- 697

Qy 508 GAFMNRIGTDLKOKLUIKGDAAHINHIIYAVLPESVQELGRAFRQNGALHM--- 563  
Db 698 -----TSBLQSKLRLADK-----LKSTEBLKKERKRRDMLGLVPMRQ 737

Db 564 -FIGKQKUTGEMALSNKLESVNLSEBOKLTIQAFSDNAEAVVLPNLQITREAA 622  
Db 738 SIIDLUKEKEPHE--LRNKLQVNDIQRKNDIQRQ---ETLGTIMP----EES 784

Qy 623 FK-----RNHLIKEKVGSSTLSQTENAPDNDGDRFSGKVVWRHN----- 664  
Db 785 AKVCLJDVTIMERQMLKOVERKAQQAQKLOGDLDRTVQVQVKOQKQHADTVSS 844

Qy 665 ---NSHMLADGERFIDPKLSSNTWDLKVKLTIEGLYST---LROPTOTPREMTT 717  
Db 845 KIELNKLUQDQEQT---OHLKSTINELKS---EKLOQISTNQORRQOLEQVTELST 896

Db 718 AGKALISKSNLRQGKQ-----KELQAFQETGVRDQKATAKAEGKALVTKATK 768

Qy 897 EVQSLYRE--IKDAKBOVSPLETTLKFOOKKEELINKNTNSKIAQ-DKLDNTEKEVN 953

Db 769 GHLLERSINKAVLAVNNSAITKANKVRLKEELDLDLTVAGKGPRAQ-ATMVOGVYLK 826

Db 954 IHGMYKMDIEN-YIQLGSKDYYTKQKETELNLYIAQSECEHKKEKNDNEDBLMRDIDTK 1012

Db 827 TPLPLIPEYYVIGLNVYFDKSKLIVALMDMSDTIGEGQDQDAYGNPILNVDNEGYHTLAVA 886

Db 1013 ---I-QERWMLQDNLTIRKRE-----ELKEVEERKQH---LK 1043

Qy 887 TLADYEGLYIK--DILNSLIDKIRAIROLPLAKTHRGLIQAFNAAAEADRLPKTPK 943

Db 1044 EMGQOMVQLOMKSEHOKLEENIDNIRNNHNLALGR-----QK 1079

Qy 944 GYLNQVNPYKQKEMKNUKVDYK-----TPIFNKALPKNEVKDGRA-AKGHNIN 992

Db 1080 GYEEWLTIFKELRPFQFDDAEKKREMMIVMRTTELVKNDLDTYKTDQATMKPHSM 1139

Qy 993 AETNNSVAVTPIRSH--QOLH---KSOSDVNLQTSKSNNFYIEL--GYVSLCLIF 1042  
Db 1140 MEEINKITIRDIWRSYRGQDIEVIEIRSDADENVSASDRRNNYVWMLKGDTALDWRG 1199

Qy 1043 LVTACK 1049  
Db 1200 RCSAQK 1206

Search completed: April 27, 2004, 11:18:08  
Job time : 28 secs